

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 897.829 Seconds  
(without alignments)  
11183.052 Million cell updates/sec

Title: US-09-513-888c-1\_COPY\_112\_456  
Perfect score: 345  
Sequence: 1 tcgagtgatcaagctgcgcaaa.....gcaggagaaagcaaatggg 345

Scoring table: IDENTITY NUC  
Gapop 10\_0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rtd.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	345	100.0	9108	9	AF123653	AF123653 Homo sapi
2	345	100.0	173264	2	AC025853	AC025853 Homo sapi
3	286	82.9	633	9	AF123654	AF123654 Homo sapi
4	286	82.9	1515	9	AF123656	AF123656 Homo sapi
5	286	82.9	1614	9	AF123655	AF123655 Homo sapi
6	286	82.9	1692	9	AF123657	AF123657 Homo sapi
7	286	82.9	1722	9	AF123658	AF123658 Homo sapi
8	286	82.9	5492	9	AF123659	AF123659 Homo sapi
9	232.8	67.5	227884	2	AC114995	AC114995 Mus muscu
10	232.8	67.5	263546	2	AC099416	AC099416 Mus muscu
11	230.8	66.9	191210	2	AC108987	AC108987 Rattus no
12	106	30.7	231	9	AF123652	AF123652 Homo sapi
13	42.4	12.3	90348	1	AF497482	AF497482 Micromono
14	39.6	11.5	1981	6	AR103043	AR103043 Sequence
15	39.6	11.5	1981	6	AR103044	AR103044 Sequence
16	39.6	11.5	1981	6	AR103045	AR103045 Sequence
17	39.6	11.5	1981	6	AR103046	AR103046 Sequence
18	39.6	11.5	1981	6	AR103047	AR103047 Sequence
19	39.6	11.5	1981	6	AR103048	AR103048 Sequence
20	39.6	11.5	1981	6	E24180	E24180 Alpha-amyla
21	39.6	11.5	1981	6	E24181	E24181 Alpha-amyla
22	39.6	11.5	1981	6	E24182	E24182 Alpha-amyla
23	39.6	11.5	1981	6	E24183	E24183 Alpha-amyla
24	39.6	11.5	1981	6	E24184	E24184 Alpha-amyla
25	39.6	11.5	1981	6	E24185	E24185 Alpha-amyla
26	39.6	11.5	3496	1	PSEMFA	D10769 Pseudomonas
27	39.6	11.5	3496	6	E03105	E03105 DNA encodin
28	39.6	11.5	177559	4	AC091401	AC091401 Sus scrof
29	39.6	11.2	324	6	E15909	E15909 gDNA encodi
30	38.6	11.2	324	6	E30020	E30020 Mutated cyt
31	38.6	11.2	873	1	DVUPCC3	D31702 Sulfate-red
32	38.6	11.2	873	6	E37765	E37765 Process for
33	38.6	11.2	873	6	E37766	E37766 Process for
34	38.6	11.2	873	6	E64601	E64601 Process for
35	38.4	11.1	6058	1	AF279141	AF279141 Rhodococc
36	38.2	11.1	10295	1	AE012032	AE012032 Xanthomon
37	37.4	10.8	196050	1	AL646058	AL646058 Ralstonia
38	37.4	10.8	230138	14	AF232689	AF232689 Rat cytom
39	37	10.7	11044	1	AE012279	AE012279 Xanthomon
40	36.8	10.7	25970	1	SC2H4	AL031514 Streptomy
41	36.8	10.7	146884	2	AF004675	AF004675 Oryza sat
42	36.6	10.6	1884	10	AF050418	AF050418 Mus muscu
43	36.6	10.6	12795	1	AE004467	AE004467 Pseudomon
44	36.4	10.6	2267	6	AR195560	AR195560 Sequence
45	36.4	10.6	4800	6	AR106490	AR106490 Sequence

ALIGNMENTS

RESULT 1  
AF123653  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE

AF123653  
Homo sapiens FEZ1 (FEZ1) gene, complete cds.  
AF123653.1 GI:4572463  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 9108)  
Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,

and its expression is altered in multiple human tumors  
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
99199287  
10097140  
REFERENCE  
Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,  
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
Direct Submission  
Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
Institute, 2335 10th street, Philadelphia, PA 19107, USA  
LOCATION/Qualifiers  
1. .9108  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
<112. .>5553  
/gene="FEZ1"  
Join(<112. 456, 1707. 2510, 4912. .>5553)  
/gene="FEZ1"  
Join(112. 456, 1707. 2510, 4912. .5553)  
/gene="FEZ1"  
/codon start=1  
/product="FEZ1"  
/protein\_id="AAD23834.1"  
/db\_xref="GI:4572464"  
/translation="MGSVSLISGHSFKHCRASQYKLRKSHHKKLNRYSDGLLRF  
GFSQSGHGKSKMGKSDFFYIKVQKARSHDPDYATLSSDGLGGQAGVDFDPST  
PKLMPFNSQLMSGKGAVRPAFKPLPRSGAILHSPESASHLPAAPDKPKQ  
ELKPLGCSALSDGRNSMSLPTHTSSYOLDPLVTPVPTSRFGSANNITGIV  
LQDSNMSLKALSFDSGSKLGHSHKADKGPSCVSPSTDECSIQLEQLLEREGA  
LQDQSRFEKLASLAYERPRRCDELEPGKNGKQAQSKQRAQVUHLQ  
VLQOQKRLQELLESMEKQDLLETKLRSYERKTSFPALESTQWEVCKSGEIV  
LKQOQKESQTVNAKASILGKALQKDKTRGLELRTQDLLEALTRQKSGEIV  
ENELQRKNEALLREKNVLELQELRAQAALARDMGPTTFPDVPALELERLR  
AELRERQGHQDMSSGFOHERLVWKEKEKVIQYQKQLQSVVYQNRLEKALQ  
LARGDSAGEPLVDLEGADIPYEDIIATEI"  
BASE COUNT 2021 a 2628 c 2504 g 1952 t 3 others  
ORIGIN  
Query Match 100.0%; Score 345; DB 9; Length 9108;  
Best Local Similarity 100.0%; Pred. No. 7,1e-78;  
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGG 60  
|||||  
Db 172 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGG 231  
QY 61 CTGCTGAGTTGGCTTCTCCAGGACTCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120  
|||||  
Db 232 CTGCTGAGTTGGCTTCTCCAGGACTCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 291  
QY 121 AAGAGCGAAGACTTCTTACATCAAGTTCAGCCAGAAAGCCGGGCTCCCATCACCA 180  
|||||  
Db 292 AAGAGCGAAGACTTCTTACATCAAGTTCAGCCAGAAAGCCGGGCTCCCATCACCA 351  
QY 181 GATTACCGGCACTGTCAGCGGGATTAGGGGGCCAGGCTGGGGTGAACCTTACCGCG 240  
|||||  
Db 352 GATTACCGGCACTGTCAGCGGGATTAGGGGGCCAGGCTGGGGTGAACCTTACCGCG 411  
QY 241 TCCACACCCCAAGCTCATGCCCTTCTCAATCAGTACAGTAAGTGTAGCGGGTCCG 300  
|||||  
Db 412 TCCACACCCCAAGCTCATGCCCTTCTCAATCAGTACAGTAAGTGTAGCGGGTCCG 471  
QY 301 TCGCAAGGTTAGTGGTGTGAAACGACGAGAAACCAAAATCGG 345  
|||||  
Db 472 TCGCAAGGTTAGTGGTGTGAAACGACGAGAAACCAAAATCGG 516

RESULT 2  
AC025853  
LOCUS

AC025853 173264 bp DNA linear HTG 17-JUN-2002

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Homo sapiens chromosome 8 clone RP11-353K12 map 8, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 3 unordered pieces.  
AC025853 GI:21431202  
AC025853.13 HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 173264)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 8, clone RP11-353K12  
Unpublished  
2 (bases 1 to 173264)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 173264)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L.,  
Boukhgalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,  
Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,  
Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N.,  
Mathews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,  
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Nicoll, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,  
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,  
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,  
Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,  
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 17, 2002 this sequence version replaced gi:21321864.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

## ----- Project Information

Center project name: L7454

Center clone name: 353\_K12

- \* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- \* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

- \* 1 39461: contig of 39461 bp in length
- \* 39462 39561: gap of 100 bp
- \* 39562- 108347: contig of 68786 bp in length
- \* 108348 108447: gap of 100 bp
- \* 108448 173264: contig of 64817 bp in length.

## FEATURES

source

Location/Qualifiers

1..173264

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-353K12"

/clone\_lib="RPC1-11 Human Male BAC"

BASE COUNT 47355 a 38864 c 39235 g 47493 t 317 others

## ORIGIN

```

Query Match      100.0%; Score 345; DB 2; Length 173264;
Best Local Similarity 100.0%; Pred. No. 6.9e-78;
Matches 345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
Db 22393 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 22452

QY 61 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120
Db 22453 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 22512

QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 180
Db 22513 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 22572

QY 181 GATTACACGGCACTGTCACGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCCG 240
Db 22573 GATTACACGGCACTGTCACGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCCG 22632

QY 241 TCCACACCCCCCAAGCTATGCCCTTCTCCATCAGTCTAGAAATGGTAAGCGGGGTGCG 300
Db 22633 TCCACACCCCCCAAGCTATGCCCTTCTCCATCAGTCTAGAAATGGTAAGCGGGGTGCG 22692

QY 301 TCGCAAGGTAAGTGGTTGGAAACGACGAGAAAGCAAAATGGG 345
Db 22693 TCGCAAGGTAAGTGGTTGGAAACGACGAGAAAGCAAAATGGG 22737

```

## RESULT 3

AF123654

LOCUS

AF123654 Homo sapiens clone E264162 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999  
 DEFINITION complete cds.

ACCESSION

AF123654

VERSION

AF123654.1

KEYWORDS

GI:4572465

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 633)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,

Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,

and its expression is altered in multiple human tumors  
 Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

## JOURNAL

MEDLINE

99199287

PUBMED

10097140

REFERENCE

2 (bases 1 to 633)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,

Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

Direct Submission

TITLE

Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer

JOURNAL

Institute, 2338 10th street, Philadelphia, PA 19107, USA

FEATURES

Location/Qualifiers

1..633

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8p22"

/clone="E264162"

1..633

/gene="FEZ1"

/note="alternatively spliced"

/codon\_start=1

/product="FEZ1"

/protein\_id="AAD23835.1"

/db\_xref="GI:4572466"

/translation="MGVSLSLIHSFHSKCRASQYKLRKSSHLKLNRYSDGLLRF

GFSODSGHKSSKMGKSEDFYIKVQKARGSHDPDYATLSSGDLGGAGVDFDPST

PPKLPFNSQNLQEMSGEKAIVPTAFKVLPRGAILHSHPESASHOLHPAPPDKPKKEQ

ELKPGLCGSAUSDSGRNSMPLTHSAGEPUEVDLEGADIPYEDIIATEI"

BASE COUNT 136 a 217 c 175 g 105 t

ORIGIN

```

Query Match      82.9%; Score 286; DB 9; Length 633;
Best Local Similarity 100.0%; Pred. No. 9.1e-63;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120
Db 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 180

QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 180
Db 181 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCA 240

QY 181 GATTACACGGCACTGTCACGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCCG 240
Db 241 GATTACACGGCACTGTCACGCGGGAATTTAGGGGCGCAGCTGGGTGGACTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTATGCCCTTCTCCATCAGTCTAGAAATGG 286
Db 301 TCCACACCCCCCAAGCTATGCCCTTCTCCATCAGTCTAGAAATGG 346

```

## RESULT 4

AF123656

LOCUS

AF123656 Homo sapiens clone D14 FEZ1 (FEZ1) mRNA, alternatively spliced,  
 DEFINITION complete cds.

ACCESSION

AF123656

VERSION

AF123656.1

KEYWORDS

GI:4572469

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1515)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,

Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
 MEDLINE 99199287  
 PUBLISHED 10097140  
 REFERENCE 2 (bases 1 to 1515)  
 AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA

## FEATURES

Source

1..1515

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8p22"

/clone="D14"

1..1515

/gene="FEZ1"

1..1515

/note="alternatively spliced"

/codon\_start=1

/product="FEZ1"

/protein\_id="AAD23837.1"

/db\_xref="GI:4572470"

/translation="MGSVSLISGHFSHKICRASYKLRKSHLKKLNRYSDGLLRF  
 GFSQGHGKSSKMGKEDFYIKVQKARGSHRPTALSSGDLGGQVDFPST  
 PKLMPFNQLEMSKGAFTAPKPLPRSGAILHSPESASHQHPADPKPEQ  
 ELKPGCLSDSNGSNSSLPTHSTSSYOLDPLVTPVGTSTRFGSAHNITOGIV  
 LQDSNMSLKALSFDSGSKLGHNSKADKPCVSPISITDECSIQLEQLKLEREGA  
 LQKQRFEEKELASLAYERPRRCRDELEGPPEKGNKLOKQASQKSORAQVHLHQ  
 VLQOQKRLQLESLMKQDLETLKLSYERKTSFGPALETQMEVCQSGEIS  
 LKQOQKESQTEVNAKSEILGLKQKTRKLEGLRLTQDLGALRTKGLLEVC  
 ENELQOSYVAMYNORLEKALQOLARGDSAGEPLEVDLEGADIPYEDIATEI"

336 a 483 c 482 g 214 t

BASE COUNT

ORIGIN

Query Match 82.9%; Score 286; DB 9; Length 1515;

Best Local Similarity 100.0%; Pred. No. 9e-63;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60

DB 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGTCACGGCAAGTCCAGCTCCAAATGGGC 120

DB 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGTCACGGCAAGTCCAGCTCCAAATGGGC 180

QY 121 AGAGCGAAGACTTCTTACATCAAGGTTCAGGCAAGTCCAGCTCCAAATGGGC 180

DB 181 AGAGCGAAGACTTCTTACATCAAGGTTCAGGCAAGTCCAGCTCCAAATGGGC 240

QY 181 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240

DB 241 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 286

DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346

RESULT 5

AF123655

LOCUS

DEFINITION

Homo sapiens clone T8D145M4 FEZ1 (FEZ1) mRNA, linear

spliced, complete cds.

AF123655

ACCESSION

AF123655.1

VERSION

GI:4572467

KEYWORDS

Homo sapiens.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1614)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors

Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)

99199287

2 (bases 1 to 1614)

Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.

Direct Submission

Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA

10097140

1..1614

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="8"

/map="8p22"

/clone="T8D145M4"

1..1614

/gene="FEZ1"

1..1614

/note="alternatively spliced"

/codon\_start=1

/product="FEZ1"

/protein\_id="AAD23836.1"

/db\_xref="GI:4572468"

367 a 500 c 522 g 225 t

BASE COUNT

ORIGIN

Query Match 82.9%; Score 286; DB 9; Length 1614;

Best Local Similarity 100.0%; Pred. No. 9e-63;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60

DB 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGTCACGGCAAGTCCAGCTCCAAATGGGC 120

DB 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGTCACGGCAAGTCCAGCTCCAAATGGGC 180

QY 121 AGAGCGAAGACTTCTTACATCAAGGTTCAGGCAAGTCCAGCTCCAAATGGGC 180

DB 181 AGAGCGAAGACTTCTTACATCAAGGTTCAGGCAAGTCCAGCTCCAAATGGGC 240

QY 181 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240

DB 241 GATTACAGGCACTGTCACGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 286

DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346

RESULT 6

AF123657



LOCUS AF123657 1692 bp mRNA linear PRI 07-APR-1999  
 DEFINITION Homo sapiens clone G3611 FEZ1 (FEZ1) mRNA, alternatively spliced,  
 complete cds.  
 ACCESSION AF123657  
 VERSION AF123657.1 GI:4572471  
 KEYWORDS  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1692)  
 AUTHORS Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H.,  
 Mori, M., Fidanza, V., Alder, H. and Croce, C. M.  
 TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,  
 and its expression is altered in multiple human tumors  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
 MEDLINE 99199287  
 PUBMED 10097140  
 REFERENCE 2 (bases 1 to 1692)  
 AUTHORS Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H.,  
 Mori, M., Fidanza, V., Alder, H. and Croce, C. M.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
 Institute, 2338 10th street, Philadelphia, PA 19107, USA  
 FEATURES  
 Location/Qualifiers  
 source  
 1. .1692  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8p22"  
 /clone="G3611"  
 1. .1692  
 /gene="FEZ1"  
 1. .1692  
 /gene="FEZ1"  
 /note="alternatively spliced"  
 /codon\_start=1  
 /product="FEZ1"  
 /protein\_id="AAD23838.1"  
 /db\_xref="GI:4572472"  
 /translation="MGSVSLISGHSFKHCRASQYKLRKSHLKLNRYSDDLRF  
 GFSQDSGHKSSKMGSEDFYIKVQKARGSHHPDYTALSSGDLGGQGVDFPST  
 PPKLPPFNSQLMSEKGAVRTAPKPLVPSRGAIIHSSPESASHQHPAPDKPEQ  
 ELKPLGCSGALSDGRNMSLPTHTSSSYQLDPLVTPVPTSRFGSAHNIQTQIV  
 LODSNMMSLKALSFDSGSKLGHKNADKPGSCVRSPISTDECSIQLEQLKLEREGA  
 LQKORSPEEKELASSLAYEERPRRCRDELEGPPEKGNKQAOKSOKRAQOVLHLQ  
 VLQQLKESQTEVNAKASEILGLKQLKTRKLEGLRTQDLSEGLTKGLELEVC  
 ENELQKKNKNEALLREKVNLLERLRAELREERQGHQDMSGQFHERLVWKEKEKVIQ  
 YOKLQSQSVAMYQNRLEKALQQLARGSDAGEPLEVDLEGADIPYDIATEI"  
 BASE COUNT 380 a 522 c 554 g 236 t

Query Match 82.9%; Score 286; DB 9; Length 1692;  
 Best Local Similarity 100.0%; Pred. No. 9e-63;  
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60  
 DB 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120  
 QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGCAAGTCCAGCTCCAAATATGGGC 120  
 DB 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGCAAGTCCAGCTCCAAATATGGGC 180  
 QY 121 AAGAGCGAAGACTTCTTCTACATCAAGGTCAGCCAGAAAGCCGGGCTCCCATCACCA 180  
 DB 181 AAGAGCGAAGACTTCTTCTACATCAAGGTCAGCCAGAAAGCCGGGCTCCCATCACCA 240  
 QY 181 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240  
 DB 241 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300

QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGG 286  
 DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGG 346  
 RESULT 7  
 AF123658 1722 bp mRNA linear PRI 07-APR-1999  
 LOCUS AF123658  
 DEFINITION Homo sapiens clone G3612 FEZ1 (FEZ1) mRNA, alternatively spliced,  
 complete cds.  
 ACCESSION AF123658  
 VERSION AF123658.1 GI:4572473  
 KEYWORDS  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1722)  
 AUTHORS Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H.,  
 Mori, M., Fidanza, V., Alder, H. and Croce, C. M.  
 TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,  
 and its expression is altered in multiple human tumors  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
 MEDLINE 99199287  
 PUBMED 10097140  
 REFERENCE 2 (bases 1 to 1722)  
 AUTHORS Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H.,  
 Mori, M., Fidanza, V., Alder, H. and Croce, C. M.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
 Institute, 2338 10th street, Philadelphia, PA 19107, USA  
 FEATURES  
 Location/Qualifiers  
 source  
 1. .1722  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8p22"  
 /clone="G3612"  
 1. .1722  
 /gene="FEZ1"  
 1. .1722  
 /gene="FEZ1"  
 /note="alternatively spliced"  
 /codon\_start=1  
 /product="FEZ1"  
 /protein\_id="AAD23839.1"  
 /db\_xref="GI:4572474"  
 /translation="MGSVSLISGHSFKHCRASQYKLRKSHLKLNRYSDDLRF  
 GFSQDSGHKSSKMGSEDFYIKVQKARGSHHPDYTALSSGDLGGQGVDFPST  
 PPKLPPFNSQLMSEKGAVRTAPKPLVPSRGAIIHSSPESASHQHPAPDKPEQ  
 ELKPLGCSGALSDGRNMSLPTHTSSSYQLDPLVTPVPTSRFGSAHNIQTQIV  
 LODSNMMSLKALSFDSGSKLGHKNADKPGSCVRSPISTDECSIQLEQLKLEREGA  
 LQKORSPEEKELASSLAYEERPRRCRDELEGPPEKGNKQAOKSOKRAQOVLHLQ  
 VLQQLKESQTEVNAKASEILGLKQLKTRKLEGLRTQDLSEGLTKGLELEVC  
 ENELQKKNKNEALLREKVNLLERLRAELREERQGHQDMSGQFHERLVWKEKEKVIQ  
 YOKLQSQSVAMYQNRLEKALQQLARGSDAGEPLEVDLEGADIPYDIATEI"  
 BASE COUNT 381 a 541 c 563 g 237 t

Query Match 82.9%; Score 286; DB 9; Length 1722;  
 Best Local Similarity 100.0%; Pred. No. 9e-63;  
 Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60  
 DB 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120  
 QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGCAAGTCCAGCTCCAAATATGGGC 120  
 DB 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGCAAGTCCAGCTCCAAATATGGGC 180



Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

Direct Submission  
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 26, 2002 this sequence version replaced gi:21535965.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L23473

Center clone name: 171\_P\_5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 225162 bases at least Q40

Consensus quality: 226535 bases at least Q30

Consensus quality: 227117 bases at least Q20

Insert size: 225000; agarose-fp

Insert size: 227484; sum-of-contigs

Quality coverage: 8.6 in Q20 bases; agarose-fp

Quality coverage: 8.6 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 5 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have

\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 301: contig of 301 bp in length

\* 302 401: gap of 100 bp

\* 402 5831: contig of 5430 bp in length

\* 5832 5931: gap of 100 bp

\* 5932 173029: contig of 167098 bp in length

\* 173030 173129: gap of 100 bp

\* 173130 210061: contig of 36932 bp in length

\* 210062 210161: gap of 100 bp

\* 210162 227884: contig of 17723 bp in length.

Location/Qualifiers

1. .227884

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="RP23-171P5"

/clone\_lib="RP23-171P5" Female Mouse BAC"

1. .301

/notes="assembly\_fragment"

clone end:SP6

vector side:left"

misc\_feature

402..5831

/notes="assembly\_fragment"

misc\_feature

5932..173029

/notes="assembly\_fragment"

misc\_feature

173130..210061

/notes="assembly\_fragment"

misc\_feature

210162..227884

/notes="assembly\_fragment"

clone end:T7

vector side:right"

BASE COUNT 65784 a 48929 c 49902 g 62967 t 402 others

ORIGIN

Query Match 67.5%; Score 232.8; DB 2; Length 227884;

Best Local Similarity 84.7%; Pred. No. 3.5e-49;

Matches 261; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60

DB 82628 TCACAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAATGATCTCAGATGG 82687

QY 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATGGGC 120

DB 82688 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAAATGGGA 82747

QY 121 AAGCGGAAGACTTCTTACATCAAGGTCAAGCAAGGCGGGGCTCCCATCACCCA 180

DB 82748 AAAAGCGAAGACTTCTTACATCAAGGTCAAGCAAGGCGGGGCTCCCATCGCCA 82807

QY 181 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGCGCAGCTGGGTGGACTTTGACCCG 240

DB 82808 GACTACACAGCCCTGTCCAGTGGGACATAGGGGTCAAGCGGAGTAGATTGATCCA 82867

QY 241 TCACACACCCCAAGCTCATGCCCTTCTCCAATCAGTAGAATGTTAAGCGGGGCTCGC 300

DB 82868 GCCACCCCAAGAGTCTATGCCCTTCTCCAATCAGTAGAATGTTAAGCGGGGCTCGC 82927

QY 301 TGGCAAGG 308

DB 82928 TGACAGGG 82935

RESULT 10

AC099416/c

LOCUS

DEFINITION

AC099416

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC099416

Mus musculus chromosome UNK clone RP23-122M11, WORKING DRAFT

SEQUENCE, 8 unordered pieces.

AC099416

AC099416.2 GI:21105058

HTG; HTGS PHASE1; HTGS\_DRAFT.

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

McPherson, J.D. and Waterston, R.H.

Unpublished

2 (bases 1 to 263546)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (14-NOV-2001) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

3 (bases 1 to 263546)

McPherson, J.D. and Waterston, R.H.

Direct Submission

Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park

Parkway, St. Louis, MO 63108, USA

On May 23, 2002 this sequence version replaced gi:16924178.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)

----- Project Information

Center project name: M\_BA0122M11

----- Summary Statistics

Sequencing vector: M13; 32%

Sequencing vector: plasmid; 68%

Chemistry: Dye-primer ET; 0% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 278310 bases at least Q40  
 Consensus quality: 282938 bases at least Q30  
 Consensus quality: 286149 bases at least Q20  
 Insert size: 250000; agarose-fp  
 Insert size: 387570; sum-of-contigs  
 Quality coverage: 25.74 in Q20 bases; agarose-fp  
 Quality coverage: 16.86 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1204: contig of 1204 bp in length  
 1205 1304: gap of unknown length  
 1305 2524: contig of 1220 bp in length  
 2525 2624: gap of unknown length  
 2625 4136: contig of 1512 bp in length  
 4137 4236: gap of unknown length  
 4237 16434: contig of 12198 bp in length  
 16435 16535: gap of unknown length  
 32106: contig of 15572 bp in length  
 32107 32206: gap of unknown length  
 32207 140224: contig of 108018 bp in length  
 140225 140324: gap of unknown length  
 140325 262693: contig of 122369 bp in length  
 262694 262793: gap of unknown length  
 262794 263546: contig of 753 bp in length.

## FEATURES

## source

1. 263546  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="UNK"  
 /clone="RP23-122M11"

misc\_feature 1. 1204  
 /note="assembly\_name:Contig127"  
 misc\_feature 1305. 2524  
 /note="assembly\_name:Contig157"  
 misc\_feature 2625. 4136  
 /note="assembly\_name:Contig360"  
 misc\_feature 4237. 16434  
 /note="assembly\_name:Contig396"  
 misc\_feature 16535. 32106  
 /note="assembly\_name:Contig397"  
 misc\_feature 32207. 140224  
 /note="assembly\_name:Contig398"  
 misc\_feature 140325. 262693  
 /note="assembly\_name:Contig399"  
 misc\_feature 262794. 263546  
 /note="assembly\_name:Contig154"  
 BASE COUNT 76628 a 56149 c 55623 g 74436 t 710 others

## ORIGIN

Query Match 67.5%; Score 232.8; DB 2; Length 263546;  
 Best Local Similarity 84.7%; Pred. No. 3.5e-49;  
 Matches 261; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
 QY 1 TCGAGTACAGTGGCAGAGTCTCCACCTCAAGAGCTCAACGGTATTCGACGGG 60  
 Db 252258 TCACAGTACAGTGGCAGAGTCTCCACCTCAAGAGCTCAACGGTATTCGAGTGG 252199  
 QY 61 CTGCTGAGTTTGGCTTCTCCAGGACTCGGTCAGGCAAGTCCAGCTCCAAATGGGC 120  
 Db 252198 CTGCTGAGTTTGGCTTCTCCAGGACTCAGGCGGTGGAAGTCAAGTTCCAAATGGGA 252139  
 QY 121 AAGAGCGAAGACTTCTTCTACATCAAGGTGAGCAGAAAGCCCGGGCTCCCATCACCCA 180

Db 252138 AAAAGCGAAGACTTCTTCTACATCAAGTCAAGAGCGCGAGGCTCCCATCGCCA 252079  
 QY 181 GATTACAGGCACATCTCCAGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTTACCCG 240  
 Db 252078 GACTACACAGCCCTGTCTCCAGTGGGACATAGGGGTGACAGCCGGAGTAGATTTTATCCA 252019  
 QY 241 TCCACACCCCAAGCTCATGCCCTTCTCCATAGTCTAGAAATGCTAAGCGGGGTCCG 300  
 Db 252018 GCCACCCCAAGCTCATGCCCTTCTCCATAGTCTAGAAATGCTAAGTGGTGGTCTCT 251959  
 QY 301 TGGCAAGG 308  
 Db 251958 TGACAGGG 251951

## RESULT 11

## AC108987

## LOCUS

AC108987 191210 bp DNA linear HTG 13-JUL-2002  
 Rattus norvegicus clone CH230-115K1, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 78 unordered pieces.

ACCESSION AC108987.3 GI:21737647

VERSION AC108987.3 HTGS PHASE1.

KEYWORDS HTG; HTGS PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 191210)

## AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Doutheite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louisaeg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Mashey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunolu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Roife, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 191210)

AUTHORS Worley, K.C.

### Query Match

Best Local Similarity 87.2%; Pred. No. 1.1e-48;		Matches 109; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
Matches 253; Conservative 0; Mismatches 37; Indels 0; Gaps 0;		QY 1 TCGCAGTACAAGCTCGCGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60	
Db 147545 TCCCACTACAAGCTCGCGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120		QY 61 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGCTCCAGGCAAGTCCAGCTCCAAA 114	
Db 147605 CTCTGAGGTTGGCTTCTCCAGGACTCCGGGCGGCAAGTCCAGCTCCAAAATGGGG 147664		Db 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAAGCTCCAGGCAAGTCCAGCTCCAAA 174	
QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGCTCAGCAGAAAGCCGGGCTCCCATCACCA 180		RESULT 13	
Db 147665 AAGAGCGAAGACTTCTTACATCAAGTCAAGCTCAGCAGAAAGCCGGGCTCCCATCACCA 147724		AF497482/c	
QY 181 GATTACACGCGACTGCCAGCGGGATTTAGGGGCGCAGCTGGGGTGAGCTTTGACCG 240		LOCUS	
Db 147725 GACTACACAGCAGTCTCCAGTGGGACATGGGGCGCAGCTGGAGTAGATTGTATCCA 147784		DEFINITION	
QY 241 TCCACACCCCGCAGCTCATGCCCTTCTCAATCAGCTAGAAATGTAAG 290		ACCESSION	
Db 147785 GCACCTCCACCAAGCTCATGCCCTTCTCCAGTCAACTGGAATGGTGA 147834		AF497482 AF505622	
RESULT 12		VERSION	
AF123652		AF497482.1 GI:22255833	
DEFINITION		KEYWORDS	
Homo sapiens clone E16T8 FE21 (FE21) mRNA, alternatively spliced, complete cds.		SOURCE	
ACCESSION		ORGANISM	
AF123652		Micromonospora echinospora.	
KEYWORDS		Micromonospora echinospora	
SOURCE		Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
ORGANISM		Micromonosporineae; Micromonosporaceae; Micromonospora.	
REFERENCE		1 (bases 1 to 90348)	
AUTHORS		Ahler, J., Shepard, E., Lomovskaya, N., Zazopoulos, E., Staffa, A.,	
TITLE		Bachmann, B.O., Huang, K., Fonstein, L., Crisny, A., Whitam, R.E.,	
JOURNAL		Farnet, C.M., and Thorson, J.S.	
REFERENCE		The calicheamicin gene cluster and its iterative type I PKS	
AUTHORS		Science (2002) In press	
TITLE		2 (bases 1 to 90348)	
JOURNAL		Zazopoulos, E.	
FEATURES		Direct Submission	
source		Submitted (19-MAR-2002) Ecopia BioSciences Inc., 7290 Frederick	
1. .90348		Banting, Montreal, QC H4S 2A1, Canada	
/organism="Micromonospora echinospora"		On Aug 15, 2002 this sequence version replaced gi:21637112.	
/strain="NRRL 15839"		Location/Qualifiers	
/sub_species="calichensis"		1. .90348	
/db_xref="taxon:1877"		/gene="calu1"	
1. .873		/gene="calu1"	
gene		/codon_start=1	
CDS		/transl_table=1	
1. .873		/product="Calu1"	
/gene="calu1"		/protein_id="AAM94759.1"	
/translation="MTLPDDPSAVAWRLVEVDFDADPEFAGLVQAMREQVPADAVHA		/db_xref="GI:22255834"	
LVVGEWGEAYERPLVPELVAAAREWALRAVFLADLYEQCEIWLNRGVDVPLAA		/translation="MTFGSHARTFAGRPVVDVVDGGLPTVDGAVSWRFLMWFDETA	
YPALVLMVVRGQDLRLPVRHRLRLEFESGGLPVAVRAGVACELPALHRLLEWL		DYKLSDTFRAAFDRFVQAAGPSVDSLVLVGMVYAAFNPAIQAQCAAAPLALRALF	
GRSDYGGDATVDDLGVLAGTGLPALRLAPCNARITADAVAAVAAAPVPRLEVLDL		LGDMPSECEVSMVMKVGVDSPLLTAYPELVLRVGGEDFAFSPVHRRLRLAVETG	
SMGLTLDGKALLAGQPLTLRLDLHHHLYSEEMAAVVAALPGVRVDVSDPQTAD		GLSRFVGNVLSDPLPALTDLELWLGASDYGGAQVSDIAPLLAGERFGLRLGLRN	
EVDGTRYRYTEVGE"		AEIADDLAEALTAARVPRLEKRLDLSLGLTGDRGAALFAGQPLKHLTLELDLHHHLYS	
880. .1827		EEMAAVVAALPGVRVDVSDPDKRADEYDGESYRYTAVSE"	
/gene="calu2"		1874. .2935	
880. .1827		/gene="calu3"	
/gene="calu2"			
/codon_start=1			
/transl_table=1			
/product="Calu2"			
/protein_id="AAM94760.1"			
/db_xref="GI:22255835"			
/translation="MTFGSHARTFAGRPVVDVVDGGLPTVDGAVSWRFLMWFDETA			
DYKLSDTFRAAFDRFVQAAGPSVDSLVLVGMVYAAFNPAIQAQCAAAPLALRALF			
LGDMPSECEVSMVMKVGVDSPLLTAYPELVLRVGGEDFAFSPVHRRLRLAVETG			
GLSRFVGNVLSDPLPALTDLELWLGASDYGGAQVSDIAPLLAGERFGLRLGLRN			
AEIADDLAEALTAARVPRLEKRLDLSLGLTGDRGAALFAGQPLKHLTLELDLHHHLYS			
EEMAAVVAALPGVRVDVSDPDKRADEYDGESYRYTAVSE"			
1874. .2935			
/gene="calu3"			
gene			
CDS			
1. .231			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosome="8"			
/map="8p22"			
/clone="E16T8"			
1. .231			
/gene="FE21"			
1. .231			
/gene="FE21"			
/note="alternatively spliced"			
/codon_start=1			
/product="FE21"			
/protein_id="AAD23833.1"			
/db_xref="GI:4572462"			
/translation="MSVSSLSIGHSFHSHKCRASQYKLRKSSHLKLNRYSDGLLRF			
GFQSDGHGKAMTRCPRASMSGSGRRRRR"			
BASE COUNT			
ORIGIN			
50 a 70 c 70 g 41 c			
Query Match			
Best Local Similarity			
30.7%; Score 106; DB 9; Length 231;			
95.6%; Pred. No. 1e-16;			

CDS

```
1874. .2935
/gene="calu3"
/codon_start=1
/transl_table=11
/product="Calu3"
/protein_id="AAM94761.1"
/db_xref="GI:22255836"
/translation="MFRVAVLDAGLPRPAVLFWADVLTGAAPPAPDALVRVDSGEDP
EDVRLRLRSAPARGELVGLADTYAGLLRGLDRVAAGGAELNDPEDVAVMCDKRRR
HARLSAAGVPAALPVGGMAEALRAAAGWRSRVFKPAHGSSAAGVIALAVAGRR
VOAVTTVEATPEGLFNSLRPRYTDQAQVAAVLDRLAPESLHVERWLPKAGLGRVVD
LRVVVQATPHAVRAARGPLTNLHGNARGDLAALRAAAGPAFAAAMATCERVAG
CFPRVQGVDMFLVGMRRHAAVEVNAFDDLLPLVLTGDRDTYAEQVQAVTDGRWER
WFAAGRALDEREVLPCGA"
2932. .3726
/gene="calu4"
2932. .3726
/gene="calu4"
/codon_start=1
/transl_table=11
/product="Calu4"
/protein_id="AAM94762.1"
/db_xref="GI:22255837"
/translation="MIGSHDLLLLTDLRYDVAAELAEARGRTPLARALPGCRWERR
HSPASFTTAAHHAFAGELPTVPTGRHERLYAAAPPGSETAGRTDVTWFDPLPTAL
AGVYHTLCGGVGFNRRSPGLGAVLPGLFABAHWEPEFGVTSPTCLDAQDLRLAEVL
PRVPAGRLPFTPLNVAALHQPNRHHLPGAQADDLASHAAALEYVDGRIDRLFTLLTGR
GRPVFAIVCADHGTAYGEDGHTGHRIGHVDVWVTVFYAHTLHPGEW"
3723. .5048
/gene="calu1"
3723. .5048
/gene="calu1"
/codon_start=1
/transl_table=11
/product="Calu1"
/protein_id="AAM94763.1"
/db_xref="GI:22255838"
/translation="MTSTDTGLGSPYQOYLVAYPHKTSYRPLRPRLPLADVWRAEAR
DALFLYVHVPFCMRCGFCNLFTRANAPAEQVTGYLRLROAEQVADALGADAGYAR
AAFGGCTPYLTADLAELFDIATTTGARLPGVPLSVETSPATPDLRLVLAHGT
TEVSIQVOSFLDAEARAAGRPORTEVEALGAIRDARIPLYNIDLLIYGIDQADTW
RESLDAALAWPEELIYPLYVRPLTGLRRHRAEWAQKLALEYQAEVLEALDAAGY
LDDYLQRRADDPDPDPDYCCQDDGMVGLCGGARSYTTLSHYFDFYAVSVTEVRV
LDDYLQRRADDPDPDPDYCCQDDGMVGLCGGARSYTTLSHYFDFYAVSVTEVRV
PELGLRLDRGTAGGLRLTLAGLSDAAGPWLTSARVDMATGYYVR"
5045. .5941
/gene="calu5"
5045. .5941
/gene="calu5"
/codon_start=1
/transl_table=11
/product="Calu5"
/protein_id="AAM94764.1"
/db_xref="GI:22255839"
/translation="WNLAILEYNGPLASCNYPGCPYCFKAKRHDPPELLRADRAALDRFA
GWATIIDVRLSVLTPMGEGLTRSWYRDAMVSLSHLSDVRVVIQTNLAARVMDLAD
ADPRVLAALVTHYPCQVDNRNRLRCARLSGLGIRYSVGVGLPEHLGAEARALRAALP
REVLVWNAAEYRDYDAEEAAWTELDPLFGFSVRPVSLLDRPCHAGETAVSLGDT
VRRCHFIAEPICNLVDGSRWRLPRPCANATCDCHIGYVHLKPLGLRDVAGGVLER
IPAAMPSARRPVDPPLRRFS"
complement(6006. .7595)
/gene="calu1"
/gene="calu1"
/codon_start=1
/transl_table=11
/product="Calu1"
/protein_id="AAM94765.1"
/db_xref="GI:22255840"
/translation="MTATMPAETPKVSAGRSRREVLQALSLGMVGMFVSIILATVVA
NALPRIADLGSQTYTWIIVTELLAMTATVPLMGWADLYSKLLIQLSUGLFAVG
SLIAGLTPNVEVLLISRIVQGVAGGNTALATIYVMAAMIPEELRGYAGIFGAVFGV
TAGPLIGGLVLDTSWLGRWCFLIGVPFTLLSIVLLQRTLHLPTVRKRKIDMGLAL
```

gene

CDS

gene

CDS

gene

CDS

gene

CDS

gene

CDS

```
LIITSGVSTLLIWSLWAGSFAWASGWTALMVVGGALLGLAVVWVESRABEPIVPLDIF
RNRTVSLTTIASVLRVGVAMGCTVFLSQFOIALGKSPVAGLMSLPMIFGLVSGSTV
AGOLITKYKRWRKRLVAGAVMTVCMLLSTIDADTGVPLLSVYNWAVLGVGVGMLMQN
LVLVAQNDVPAHELGAATSVLTFFRSMGCTTGVSGALGAVLANRVATLSREKLGPLAGG
ATAGGHAVPDLSALPEPVLRIIQDVIYGTATAEFLIIGAPFALLAVVVVLFIREKPLH
TISGEERAREEAAAGMPCAH"
complement(7665. .8285)
/gene="calu1"
complement(7665. .8285)
/gene="calu1"
/codon_start=1
/transl_table=11
/product="Calu1"
/protein_id="AAM94766.1"
/db_xref="GI:22255841"
/translation="MHTGETSAPGRRDRKRTRAAALTGAALRLVAERGLAQVTVEE
ISEAGVSPRTFFNFYACKDDALVGDHAGDAARLVARLAAVPEPVLTLRAALADEV
IDEVTERELWLRMEVTRNRPVLPLRVAGNVETERAIEVVAARVGVDPDHGYPAL
VTAVAGAAVRTAMTRWAAHPDGPDLADLVEAFAAALAGLPQDPDT"
8447. .9382
/gene="calu6"
8447. .9382
/gene="calu6"
/codon_start=1
/transl_table=11
/product="Calu6"
/protein_id="AAM94767.1"
/db_xref="GI:22255842"
/translation="MTVTEQGAATAVPRATPTGRGRIVLGATVAGVLISVLSWSYEFVD
HVIDSVANMLGRDKATAIGTLAGLVFAFVPSGLAGTFTACTNMAASIGPMSQAG
AGGATRSALRSARLPVAMTAGWAAVAVGVGVLDGLPQLSTASVGGLPVRLVQ
SMWVFLIGLAFVVLGALGVLRPFAHRTARVVVLGALVGGFLIGRPYPLFNKLF
HWAVETGNPLYGAGAFVLSIGNAVVTALFALIVLITGRFVSVLAASPSRTAALTG
GMLIALGVPMVYVDVRLPAMPFDYGVFFTMFYNDR"
complement(9415. .9987)
/gene="calu1"
complement(9415. .9987)
/gene="calu1"
/codon_start=1
/transl_table=11
/product="Calu1"
/protein_id="AAM94768.1"
/db_xref="GI:22255843"
/translation="MDIRALTIEGSWLVTPOQHDDRGFLFLEWYQADRFTEAVGRPF
VAQSNVSGRGVIRGHLVDVPPGQKVIITCVSAGVLDVVDLVGVSGPTGEGWGR
LDDVDRRAVHLDAGLGHAFALTEPATVYVILCSSPYDPAARTVNPDLDELGIAPAG
PRRUSPRDATAQNAQAALAGVLTPTVYPRG"
10158. .11081
/gene="calu2"
10158. .11081
/gene="calu2"
/codon_start=1
/transl_table=11
/product="Calu2"
/protein_id="AAM94769.1"
/db_xref="GI:22255844"
/translation="MSVPTPLVLTGAGLLGRELGELLAGSGDRHTTLDPSALDITD
AAAYRHVAGCHDVVNAAAWTDVGAESDEKAGERAVNGGLPAGGYVVRTGWLIGHL
SSDVFSGETTPMPEDAPPALNAYGRSLKLAGERAVRGLLPAGGYVVRTGWLIGHL
PNFVRVLRRAAETEFVDVDDQHQPTWBARLGRVLAAGLAAAGAPAGIYHGTA
AGRTTWYGFARTVFALHGLDPDRVPTGSGTTLAARPARRRPRIIVLGHGWARAGLGPL
PPRRMTEAAVAEERSRAGPAGQDTRRSSA"
11078. .12109
/gene="calu3"
11078. .12109
/gene="calu3"
/codon_start=1
```

gene

CDS

gene

CDS

gene

CDS

gene

CDS

Query Match 12.3%; Score 42.4; DB 1; Length 90348;  
Best Local Similarity 49.5%; Pred. No. 1.8;  
Matches 109; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy

12 GCTGCGCAAGTCCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGCTCTCAGGTT 71

Db 46437 GCCCGCCTGACCCCAAGTCAAGCGCTCGGCCAGCACTCGTTCAAGGGCATGGCGT 46378  
QY 72 TGGCTTCTCCAGGACTCCGGTCAAGGCTAGTTCAGTCCAAATGGGCAAGCGAGA 131  
Db 46377 GGACTTCGGCGACATCGAGCGGACGGCATGTTTCGACCTGTACGTGCGCAACATCACCAC 46318  
QY 132 CTTCTTCTACATCAAGTTCAGCCAGAAAGCCCGGGCTCCCATCACCCAGATTACACGCG 191  
Db 46317 CTCCTTCGGCATCCAGGAGAGCAACTTCGCCCTTCGTCAACACCGCGCGCACCGCGCG 46258  
QY 192 ACTGTCACGCGGGGATTTAGGGGGCCAGGCTGGGTGGAC 231  
Db 46257 GCTGCGCGCGGCTCTGGCGCGGAGGCGCGCTGGCAC 46218

RESULT 14  
AR103043  
LOCUS AR103043 1981 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 3 from patent US 6087147.  
ACCESSION AR103043  
VERSION AR103043.1 GI:12814631  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS Ito,Y.  
TITLE .alpha.-amylase gene having ability for highly producing  
maltopentaose, vector containing said gene and transformant  
JOURNAL Patent: US 6087147-A 3 11-JUL-2000;  
FEATURES Location/Qualifiers  
source 1..1981  
BASE COUNT 342 a 678 c 678 g 283 t  
ORIGIN

Query Match 11.5%; Score 39.6; DB 6; Length 1981;  
Best Local Similarity 52.4%; Pred. No. 9.8; Mismatches 79; Indels 0; Gaps 0;  
Matches 87; Conservative 0;  
QY 3 GCAGTACAAGCTGCGCAAGTCTCCCACTCAAGAAGCTCAACCGGTATTCGACGGGCT 62  
Db 792 GCAGGTACGGTGACCGAGTTCGCTACGCAAGGAGCTCTACGGCAAGTTCGCGCGGG 851  
QY 63 GCTGAGTTTGGCTTCTCCAGGACTCCGGTACGGCAAGTTCAGCTCCAAATGGGCAA 122  
Db 852 CGCAAGCTGGCGACCTGCAGACCTTCGGCCCCAGCTGGAACCTGATGCCCGCAGCAA 911  
QY 123 GAGCGAAGACTTCTTACATCAAGTTCAGCCAGAAAGCCCGGGG 168  
Db 912 GGCCATCGCTTTCGTGCAACCAACAGCAGCGCGCGCCACGGC 957

RESULT 15  
AR103044  
LOCUS AR103044 1981 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 4 from patent US 6087147.  
ACCESSION AR103044  
VERSION AR103044.1 GI:12814632  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS Ito,Y.  
TITLE .alpha.-amylase gene having ability for highly producing  
maltopentaose, vector containing said gene and transformant  
JOURNAL Patent: US 6087147-A 4 11-JUL-2000;  
FEATURES Location/Qualifiers  
source 1..1981  
BASE COUNT 343 a 679 c 678 g 281 t

ORIGIN  
Query Match 11.5%; Score 39.6; DB 6; Length 1981;  
Best Local Similarity 52.4%; Pred. No. 9.8; Mismatches 79; Indels 0; Gaps 0;  
Matches 87; Conservative 0;  
QY 3 GCAGTACAAGCTGCGCAAGTCTCCCACTCAAGAAGCTCAACCGGTATTCGACGGGCT 62  
Db 792 GCAGGTACGGTGACCGAGTTCGCTACGCAAGGAGCTCTACGGCAAGTTCGCGCGGG 851  
QY 63 GCTGAGTTTGGCTTCTCCAGGACTCCGGTACGGCAAGTTCAGCTCCAAATGGGCAA 122  
Db 852 CGCAAGCTGGCGACCTGCAGACCTTCGGCCCCAGCTGGAACCTGATGCCCGCAGCAA 911  
QY 123 GAGCGAAGACTTCTTACATCAAGTTCAGCCAGAAAGCCCGGGG 168  
Db 912 GGCCATCGCTTTCGTGCAACCAACAGCAGCGCGCGCCACGGC 957

Search completed: June 15, 2003, 02:18:31  
Job time : 901.829 secs





PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -

Claim 2; Fig 5A; 255pp; English.

The present sequence comprises the human FEZ1 gene. FEZ1 is a tumour suppressor gene, located at chromosome location 9p22. Decreased or no expression of FEZ1 is detected in a variety of cancer cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts with tubulin, with microtubules, and with protein FE1-gamma. Post-translational phosphorylation and dephosphorylation modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are useful for inducing cells to proliferate. Compounds which modulate FEZ1 association with tubulin are useful for alleviating tubulin hyper- or hypo- polymerisation disorders, such as those associated with aberrant initiation of mitosis, modulation of the initiation and rate of cell proliferation and cell growth, modulation of cell shape, cell rigidity, cell motility, rate and stage of cellular DNA replication, intracellular distribution of organelles, metastatic potential of cell and cellular transformation from a non-cancerous to cancerous phenotype. Compounds which modulate FEZ1 binding and phosphorylation are also useful for alleviating a disorder, such as tumorigenesis, tumour survival, growth and metastasis.

Sequence 9048 BP; 2011 A; 2605 C; 2490 G; 1939 T; 3 other;

Query Match	100.0%;	Score	345;	DB	21;	Length	9048;
Best Local Similarity	100.0%;	Pred. No.	3.7e-88;				
Matches	345;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps							
Qy	1	TCGCGAGTACAAGCTCGCGCAAGTCTCCACACTCAAGAAGCTCAACCGGTATTC	CGACGG	60			
Db	112	TCGCGAGTACAAGCTCGCGCAAGTCTCCACACTCAAGAAGCTCAACCGGTATTC	CGACGG	171			
Qy	61	CTGCTGAGGTTTGGCTTCTCCACAGACTCTCCGGTCA	CGCGAAGTCCAGCTCCAAAAATGGGC	120			
Db	172	CTGCTGAGGTTTGGCTTCTCCACAGACTCTCCGGTCA	CGCGAAGTCCAGCTCCAAAAATGGGC	231			
Qy	121	AAGAGCGAAGACTTCTTTATCATCAAGGTCA	GCAGAAAGCCCGGGGTCCCATCAACCA	180			
Db	232	AAGAGCGAAGACTTCTTTATCATCAAGGTCA	GCAGAAAGCCCGGGGTCCCATCAACCA	291			
Qy	181	GATTACACGGCACTGCTCAGCGGGGATTTAGGGGGCC	CAGGCTGGGTGGACTTTGACCCG	240			
Db	292	GATTACACGGCACTGCTCAGCGGGGATTTAGGGGGCC	CAGGCTGGGTGGACTTTGACCCG	351			
Qy	241	TCCACACCCCCCAAGCTCATGCCCTTCTCCAAATCAGCT	GAGAAATGGTAAGCGGGGGTTCGC	300			
Db	352	TCCACACCCCCCAAGCTCATGCCCTTCTCCAAATCAGCT	GAGAAATGGTAAGCGGGGGTTCGC	411			
Qy	301	TGGCAAGGGTAAAGTGGGTGGAAACGCAAGGAGAAG	CAAAATGGG	345			
Db	412	TGSCAAGGGTAAAGTGGGTGGAAACGCAAGGAGAAG	CAAAATGGG	456			

## RESULT 2

AAA64511

ID · AAA64511 standard; cDNA; 633 BP.

XX

AC AAA64511;

XX

DT 02-JAN-2001 (first entry)

XX

DE Nucleotide sequence of truncated FEZ1 transcript E264162.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
 XX tumour proliferation; tubulin; microtubule; protein Epi-gamma;  
 KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
 KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
 KW tumorigenesis; tumour survival; metastasis; ss.  
 XX  
 XX  
 OS Homo sapiens.

```

RESULT 3
AAA64513
ID AAA64513 standard; cDNA; 1512 BP.
XX
AC AAA64513;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of truncated FEZ1 transcript D14.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein Efi-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 1..1512
FT FT /*tag= a
FT FT /product= "truncated FEZ1"
XX
PN WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Ishii H;
XX
DR WPI; 2000-558396/51.
DR P-PSDB; AAB08720.
XX
PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
PS Disclosure; Fig 5F; 255pp; English.
XX
CC The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC Efi-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 1512 BP; 335 A; 483 C; 481 G; 213 T; 0 other;
XX
Query Match 82.9%; Score 286; DB 21; Length 1512;
Best Local Similarity. 100.0%; Pred. No. 1.4e-71;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TCGCAGTACAAGCTGGCGAAGTCTCCCACTCAAGAGCTCAACCGGTATTCCGACGGG 60

```

QY

CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 1614 BP; 367 A; 500 C; 522 G; 225 T; 0 other;  
Query Match 82.9%; Score 286; DB 21; Length 1614;  
Best Local Similarity 100.0%; Pred. No. 1.4e-71;  
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCGCAGTACAAGTGGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 60  
DB 61 TCGCAGTACAAGTGGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120  
QY 61 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTTCAAAATGGGC 120  
DB 121 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTTCAAAATGGGC 180  
QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 180  
DB 181 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 240  
QY 181 GATTACACGGCACTGTCACGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 240  
DB 241 GATTACACGGCACTGTCACGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 300  
QY 241 TCCACACCCCGCAAGCTCATGTCCTTCTCAATCAGCTAGAAATGG 286  
DB 301 TCCACACCCCGCAAGCTCATGTCCTTCTCAATCAGCTAGAAATGG 346  
RESULT 5  
ID AAA64515 standard; cDNA; 1722 BP.  
XX  
AC AAA64515;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Nucleotide sequence of truncated FEZ1 transcript G3612.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein Efi-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1722  
FT FT /\*tag= a  
FT FT /product= "truncated FEZ1"  
XX  
PN W0200050565-A2.  
XX  
PD 31-AUG-2000.  
XX  
PF 25-FEB-2000; 2000WO-US04950.  
XX  
PR 25-FEB-1999; 99US-0121537.  
XX  
PA (UJJE-) UNIV JEFFERSON THOMAS.  
XX  
PI Croce CM, Ishii H;  
XX WPI; 2000-558396/51.  
DR

DR P-PSDB; AAB08722.  
XX  
PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
PS Disclosure; Fig 5H; 255pp; English.  
XX  
CC The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC Efi-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;  
Query Match 82.9%; Score 286; DB 21; Length 1722;  
Best Local Similarity 100.0%; Pred. No. 1.4e-71;  
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCGCAGTACAAGTGGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 60  
DB 61 TCGCAGTACAAGTGGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120  
QY 61 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTTCAAAATGGGC 120  
DB 121 CTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTTCAAAATGGGC 180  
QY 121 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 180  
DB 181 AAGAGCGAAGACTTCTTACATCAAGTCAAGGCAAGGCGGGCTCCCATCACCCA 240  
QY 181 GATTACACGGCACTGTCACGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 240  
DB 241 GATTACACGGCACTGTCACGCGGGGATTTAGGGGCGCAGGCTGGGTGGACTTTGACCCG 300  
QY 241 TCCACACCCCGCAAGCTCATGTCCTTCTCAATCAGCTAGAAATGG 286  
DB 301 TCCACACCCCGCAAGCTCATGTCCTTCTCAATCAGCTAGAAATGG 346  
RESULT 6  
ID AAA64509 standard; cDNA; 1791 BP.  
XX  
AC AAA64509;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE cDNA sequence encoding a human FEZ1 polypeptide.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein Efi-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX  
OS Homo sapiens.  
XX

```

FH Key Location/Qualifiers
FT CDS 1..1791
FT FT /*tag= a
FT FT /product= "FEZ1"
XX
XX WO2000050565-A2.
XX
XX 31-AUG-2000.
XX
XX 25-FEB-2000; 2000WO-US04950.
XX
XX 25-FEB-1999; 99US-0121537.
XX
XX (UYJE-) UNIV JEFFERSON THOMAS.
XX
XX Croce CM, Ishii H;
XX
XX WPI; 2000-558396/51.
XX
XX P-PSDB; AAB08715.
XX
XX New polynucleotide homologous with a portion of one strand of the human
XX FEZ1 gene, useful for alleviating abnormal cell proliferation such as
XX cancer -
XX
XX Claim 7; Fig 5I; 255pp; English.
XX
XX The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a
XX tumour suppressor gene, located at chromosome location 8p22. Decreased
XX or no expression of FEZ1 is detected in a variety of cancer cells.
XX Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
XX also interacts with tubulin, with microtubules, and with protein
XX BFI-gamma. Post-translational phosphorylation and dephosphorylation
XX modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
XX expression are useful for inducing cells to proliferate. Compounds
XX which modulate FEZ1 association with tubulin are useful for alleviating
XX tubulin hyper- or hypo- polymerisation disorders, such as those
XX associated with aberrant initiation of mitosis, modulation of the
XX initiation and rate of cell proliferation and cell growth, modulation of
XX cell shape, cell rigidity, cell motility, rate and stage of cellular
XX DNA replication, intracellular distribution of organelles, metastatic
XX potential of cell and cellular transformation from a non-cancerous to
XX cancerous phenotype. Compounds which modulate FEZ1 binding and
XX phosphorylation are also useful for alleviating a disorder, such as
XX tumorigenesis, tumour survival, growth and metastasis.
XX
XX Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;
XX
XX
XX Query Match 82.9%; Score 286; DB 21; Length 1791;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-71;
XX Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TCGCAGTACAAGTCGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60
DB 61 TCGCAGTACAAGTCGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120
QY 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGC 120
DB 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGC 180
QY 121 AGAGCGGAAGACTTCTTACATCAAGGTTCAGCAAGGTCGAGTCCAGTCCAAAATGGGC 180
DB 181 AGAGCGGAAGACTTCTTACATCAAGGTTCAGCAAGGTCGAGTCCAGTCCAAAATGGGC 240
QY 181 GATTACACGCACTGTCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 240
DB 241 GATTACACGCACTGTCAGCGGGGATTTAGGGGGCCAGCTGGGGTGGACTTTGACCCG 300
QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 286
DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346
XX
XX RESULT 7

```

AAA64508

ID AAA64508 standard; cDNA; 5492 BP.

XX

AC AAA64508;

XX

DT 02-JAN-2001 (first entry)

XX

DE cDNA sequence of the wild type human FEZ1 gene.

XX

Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
tumour proliferation; tubulin; microtubule; protein BFI-gamma;  
tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
tumorigenesis; tumour survival; metastasis; ss.

XX

OS Homo sapiens.

XX

XX WO2000050565-A2.

XX

XX 31-AUG-2000.

XX

XX 25-FEB-2000; 2000WO-US04950.

XX

XX 25-FEB-1999; 99US-0121537.

XX

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX

XX Croce CM, Ishii H;

XX

XX WPI; 2000-558396/51.

XX

New polynucleotide homologous with a portion of one strand of the human  
FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
cancer -

XX

XX Example 2; Fig 5B; 255pp; English.

XX

The present sequence represents the cDNA sequence of the human FEZ1 gene.  
FEZ1 is a tumour suppressor gene, located at chromosome location 8p22.  
Decreased or no expression of FEZ1 is detected in a variety of cancer  
cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
also interacts with tubulin, with microtubules, and with protein  
BFI-gamma. Post-translational phosphorylation and dephosphorylation  
modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
expression are useful for inducing cells to proliferate. Compounds  
which modulate FEZ1 association with tubulin are useful for alleviating  
tubulin hyper- or hypo- polymerisation disorders, such as those  
associated with aberrant initiation of mitosis, modulation of the  
initiation and rate of cell proliferation and cell growth, modulation of  
cell shape, cell rigidity, cell motility, rate and stage of cellular  
DNA replication, intracellular distribution of organelles, metastatic  
potential of cell and cellular transformation from a non-cancerous to  
cancerous phenotype. Compounds which modulate FEZ1 binding and  
phosphorylation are also useful for alleviating a disorder, such as  
tumorigenesis, tumour survival, growth and metastasis.

XX

SQ Sequence 5492 BP; 1137 A; 1704 C; 1565 G; 1086 T; 0 other;

XX

Query Match 82.9%; Score 286; DB 21; Length 5492;

XX

Best Local Similarity 100.0%; Pred. No. 2e-71;

XX

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 1 TCGCAGTACAAGTCGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60

DB

172 TCGCAGTACAAGTCGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 231

QY

61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGTCCAAAATGGGC 120

DB

232 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGTCCAAAATGGGC 291

QY

121 AGAGCGGAAGACTTCTTACATCAAGGTTCAGCAAGGTCGAGTCCAGTCCAAAATGGGC 180

DB

292 AGAGCGGAAGACTTCTTACATCAAGGTTCAGCAAGGTCGAGTCCAGTCCAAAATGGGC 351

QY 181 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGAGCTTTGACCCG 240  
DB 352 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGAGCTTTGACCCG 411  
QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 286  
DB 412 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 457

RESULT 8  
AA64514  
ID AAA64514 standard; cDNA; 1692 BP.  
XX  
AC AAA64514;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Nucleotide sequence of truncated FEZ1 transcript G3611.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EFL-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX  
OS Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT CDS 1..1692  
FT /tag= a  
FT /product= "truncated FEZ1"  
XX  
XX WO2000050565-A2.  
XX  
XX 31-AUG-2000.  
XX  
XX 25-FEB-2000; 2000WO-US04950.  
XX  
XX 25-FEB-1999; 99US-0121537.  
XX  
XX (UYJE-) UNIV JEFFERSON THOMAS.  
XX  
XX Croce CM, Ishii H;  
PI WPI; 2000-558396/51.  
DR P-PSDB; AAB08721.  
XX  
XX New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
XX Disclosure; Fig 5G; 255pp; English.

XX  
XX The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EFL-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.

XX  
SQ Sequence 1692 BP; 380 A; 523 C; 553 G; 236 T; 0 other;  
Query Match 82.4%; Score 284.4; DB 21; Length 1692;  
Best Local Similarity 99.7%; Pred. No. 4e-71;  
Matches 285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTTCCACGGG 60  
DB 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTTCCACGGG 120  
QY 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCCGTACGGCAAGTCCAGCTCCAAAATGGGC 120  
DB 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCCGTACGGCAAGTCCAGCTCCAAAATGGGC 180  
QY 121 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGCCAGCAAGCCGGGCTCCCATCACCA 180  
DB 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGCCAGCAAGCCGGGCTCCCATCACCA 240  
QY 181 GATTACACGGCACTGTCCAGGGGATTTAGGGGCCAGGCTGGGGTGAGCTTTGACCCG 240  
DB 241 GATTACACGGCACTGTCCAGGGGATTTAGGGGCCAGGCTGGGGTGAGCTTTGACCCG 300  
QY 241 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 286  
DB 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGG 346

RESULT 9  
AAS30637  
ID AAS30637 standard; DNA; 2411 BP.  
XX  
AC AAS30637;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE DNA encoding novel lung cancer antigen, Seq ID No 89.  
XX  
KW Human; lung cancer; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;  
KW ophthalmological; autoimmune disease; rheumatoid arthritis; neoplasia;  
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; skin aging;  
KW ocular disorder; wound healing; organ transplantation; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200155300-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01238.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465565/50.  
XX  
XX Isolated nucleic acid molecule encoding a lung cancer antigen is used  
PT in preventing, treating or ameliorating a medical condition -  
XX  
XX Disclosure; SEQ ID No 89; 475pp; English.  
XX  
XX The invention relates to novel isolated lung cancer antigen  
CC polynucleotides (I) and polypeptides (II). (I) and (II) are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are  
CC also used in diagnosing a pathological condition or susceptibility to a  
CC pathological condition, in particular, lung cancer. The antibodies to

CC (II) can also be used in alleviating symptoms associated with the  
 CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
 CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
 CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
 CC ocular disorders e.g. corneal infection. The polypeptides can also be  
 CC used to aid wound healing and epithelial cell proliferation, to prevent  
 CC skin aging due to sunburn, to maintain organs before transplantation, for  
 CC supporting cell culture of primary tissues, to regenerate tissues and in  
 CC chemotaxis. AAS30580-AAS30685 represent novel human lung cancer antigen  
 CC coding sequences, PCR primers and related sequences of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at: ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 2411 BP; 558 A; 681 C; 733 G; 439 T; 0 other;

Query Match 33.3%; Score 115; DB 22; Length 2411;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-23;  
 Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 231 CTTTGACCGTCCACACCCCAAGCTCATGCCCTTCTCAATCAGCTAGAAATGGTAAG 290  
 Db 1 CTTTGACCGTCCACACCCCAAGCTCATGCCCTTCTCAATCAGCTAGAAATGGTAAG 60  
 QY 291 CGGGGGTCTGCGCAAGGTAAGTGGTTGGAAACGACGAGAAAGCAAAATGGG 345  
 Db 61 CGGGGGTCTGCGCAAGGTAAGTGGTTGGAAACGACGAGAAAGCAAAATGGG 115

# RESULT 10

AAS28699

ID AAS28699 standard; DNA; 2411 BP.

AC AAS28699;

DT 07-NOV-2001 (first entry)

DE Genomic sequence #539 encoding for novel human respiratory antigen.

KW Human; respiratory antigen; respiratory disorder; throat disorder;  
 KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;  
 KW anti allergic; anti asthmatic; anti inflammatory; olfactory;  
 KW respiratory active; ds.

XX Homo sapiens.

XX WO200155448-A1.

PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01333.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0232403.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 21-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 29-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 08-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 06-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-476224/51.  
XX  
XX Isolated polypeptide for treating, preventing and/or diagnosing  
XX disorders related to the respiratory system including respiratory  
XX cancers and also for testing and detection e.g. diagnosis -  
XX  
XX Disclosure; SED ID No 1133; 546pp; English.  
XX  
XX The present invention relates to the isolation of novel human  
XX respiratory antigens (AAU17685-AAU17975), and cDNA and genomic  
XX sequences encoding for these polypeptides. The sequences of the  
XX invention are useful for preventing, treating and/or diagnosing  
XX disorders related to the respiratory system including throat  
XX disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),  
XX lung disorders e.g. pneumonia, allergic disorders e.g. asthma,  
XX pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of  
XX the respiratory tissues e.g. lung cancer. The polynucleotide sequences  
XX of the invention are useful in gene therapy and antisense therapy.  
XX AAS28161-AAS28764 represent genomic sequences encoding for novel  
XX human respiratory antigens.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 2411 BP; 558 A; 581 C; 733 G; 439 T; 0 other;  
XX  
XX Query Match 33.3%; Score 115; DB 22; Length 2411;  
XX Best Local Similarity 100.0%; Pred. NO. 7.7e-23;  
XX Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX 231 CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCCCAATCAGCTAGAAATGGTAAG 290  
XX 1 CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCCCAATCAGCTAGAAATGGTAAG 60  
XX  
XX 291 CGGGGTCGCTGGCAAGGTAAGTGGGTGAAACGAGGAGAAAGCAAAATGGG 345  
XX 61 CGGGGTCGCTGGCAAGGTAAGTGGGTGAAACGAGGAGAAAGCAAAATGGG 115  
XX  
XX RESULT 11  
XX AAA64510  
XX ID AAA64510 standard; cDNA; 404 BP.  
XX  
XX AC AAA64510;  
XX  
XX 02-JAN-2001 (first entry)  
XX  
XX Nucleotide sequence of truncated FEZ1 transcript E1678.  
XX  
XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
XX tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
XX tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
XX cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
XX tumorigenesis; tumour survival; metastasis; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..231  
XX /tag= a  
XX /product= "truncated FEZ1"  
XX  
XX WO2000050565-A2.  
XX  
XX 31-AUG-2000.  
XX  
XX 25-FEB-2000; 2000WO-US04950.  
XX  
XX 25-FEB-1999; 99US-0121537.  
XX  
XX (UYJE-) UNIV JEFFERSON THOMAS.  
XX  
XX Croce CM, Ishii H;  
XX  
XX WPI; 2000-558396/51.  
XX P-PSDB; AAB08717.  
XX  
XX New polynucleotide homologous with a portion of one strand of the human  
XX FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
XX cancer -  
XX  
XX Disclosure; Fig 5C; 255pp; English.  
XX  
XX The present sequence encodes a truncated human FEZ1 polypeptide. The  
XX suppressor gene, located at chromosome location 8p22. Decreased  
XX or no expression of FEZ1 is detected in a variety of cancer cells.  
XX Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
XX also interacts with tubulin, with microtubules, and with protein  
XX EPI-gamma. Post-translational phosphorylation and dephosphorylation  
XX modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
XX expression are useful for inducing cells to proliferate. Compounds  
XX which modulate FEZ1 association with tubulin are useful for alleviating  
XX tubulin hyper- or hypo- polymerisation disorders, such as those  
XX associated with aberrant initiation of mitosis, modulation of the



CC initiation and rate of cell proliferation and cell growth, modulation of  
 CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
 CC DNA replication, intracellular distribution of organelles, metastatic  
 CC potential of cell and cellular transformation from a non-cancerous to  
 CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
 CC phosphorylation are also useful for alleviating a disorder, such as  
 CC tumorigenesis, tumour survival, growth and metastasis.  
 XX  
 SQ Sequence 404 BP; 92 A; 119 C; 127 G; 66 T; 0 other;

Query Match 30.7%; Score 106; DB 21; Length 404;  
 Best Local Similarity 95.6%; Pred. No. 1.7e-20;  
 Matches 109; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCGCAGTACAAGCTCGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60  
 DB 61 TCGCAGTACAAGCTCGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120

QY 61 CTGCTGAGGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGCTCCAAA 114  
 DB 121 CTGCTGAGGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGCTCCAAA 174

RESULT 12  
 AAX37293  
 ID AAX37293 standard; DNA; 1981 BP.  
 AC AAX37293;  
 XX  
 XX 27-JUL-1999 (first entry)  
 XX  
 XX Pseudomonas alpha-amylase variant encoding DNA (FERM BP-6116).  
 DE  
 XX Pseudomonas; maltopentaose; alpha-amylase; enzyme; variant; ss.  
 KW  
 XX Pseudomonas sp.  
 OS  
 OS Synthetic.  
 XX  
 XX JP11123081-A.  
 PN  
 XX 11-MAY-1999.  
 PD  
 XX 21-OCT-1997; 97JP-0305071.  
 PF  
 XX 21-OCT-1997; 97JP-0305071.  
 PR  
 XX (NORQ) NORINSUISANSHO SHOKUHIN SOGO.  
 PA  
 XX WPI; 1999-340517/29.  
 DR P-PSDB; AAY17904.  
 XX  
 XX New alpha-amylase gene - useful for production of malto-pentaose  
 PT  
 XX Claim 4; Page 6-8; 23pp; Japanese.  
 PS  
 XX Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase  
 CC variants (AAY17904-Y17909). The variants are derived from the  
 CC Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.  
 CC et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants  
 CC have a replacement of the 57th or the 130th amino acid residue in the  
 CC sequence as compared to the wild-type. The sequences are deposited under  
 CC the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The  
 CC variants can be used to transform plants to produce maltopentaose. The  
 CC enzymes prepared from the alpha-amylase variants produce maltopentaose  
 CC in a high yield.  
 XX  
 SQ Sequence 1981 BP; 342 A; 678 C; 678 G; 283 T; 0 other;

Query Match 11.5%; Score 39.6; DB 20; Length 1981;  
 Best Local Similarity 52.4%; Pred. No. 0.22;  
 Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 3 GCAGTACAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62  
 DB 792 GCAGTACAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 851

QY 63 GCTGAGGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGCAA 122  
 DB 852 CGGCAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 911

QY 123 GAGCGAAGACTTCTTCTACATCAAGGTCAGCGCAAGTCCAGCTCCAAAATGGGCAA 168  
 DB 912 GGCCATCGCTTTCGTGACCAACCAAGCAAGCGCGGCGGCACGGC 957

DB 792 GCAGTACAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 851  
 QY 63 GCTGAGGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGCAA 122  
 DB 852 CGGCAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 911

QY 123 GAGCGAAGACTTCTTCTACATCAAGGTCAGCGCAAGTCCAGCTCCAAAATGGGCT 168  
 DB 912 GGCCATCGCTTTCGTGACCAACCAAGCAAGCGCGGCGGCACGGC 957

RESULT 13  
 AAX37294  
 ID AAX37294 standard; DNA; 1981 BP.  
 AC AAX37294;  
 XX  
 XX 27-JUL-1999 (first entry)  
 XX  
 XX Pseudomonas alpha-amylase variant encoding DNA (FERM BP-6119).  
 DE  
 XX Pseudomonas; maltopentaose; alpha-amylase; enzyme; variant; ss.  
 KW  
 XX Pseudomonas sp.  
 OS  
 OS Synthetic.  
 XX  
 XX JP11123081-A.  
 PN  
 XX 11-MAY-1999.  
 PD  
 XX 21-OCT-1997; 97JP-0305071.  
 PF  
 XX 21-OCT-1997; 97JP-0305071.  
 PR  
 XX (NORQ) NORINSUISANSHO SHOKUHIN SOGO.  
 PA  
 XX WPI; 1999-340517/29.  
 DR P-PSDB; AAY17905.  
 XX  
 XX New alpha-amylase gene - useful for production of malto-pentaose  
 PT  
 XX Claim 5; Page 9-11; 23pp; Japanese.  
 PS  
 XX Sequences AAX37293-298 represent nucleic acids encoding alpha-amylase  
 CC variants (AAY17904-Y17909). The variants are derived from the  
 CC Pseudomonas sp. KO-8940-derived maltopentaose-forming amylase (Shida, O.  
 CC et al, Biosci. Biotech. Biochem. Vol. 56, 76-80, 1992). The variants  
 CC have a replacement of the 57th or the 130th amino acid residue in the  
 CC sequence as compared to the wild-type. The sequences are deposited under  
 CC the accession numbers FERM BP-6116, 6119, 6115, 6117, 6118 or 6114. The  
 CC variants can be used to transform plants to produce maltopentaose. The  
 CC enzymes prepared from the alpha-amylase variants produce maltopentaose  
 CC in a high yield.  
 XX  
 SQ Sequence 1981 BP; 343 A; 679 C; 678 G; 281 T; 0 other;

Query Match 11.5%; Score 39.6; DB 20; Length 1981;  
 Best Local Similarity 52.4%; Pred. No. 0.22;  
 Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 3 GCAGTACAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62  
 DB 792 GCAGTACAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 851

QY 63 GCTGAGGTTGGCTTCTCCAGAGCTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGCAA 122  
 DB 852 CGGCAAGCTGGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 911

QY 123 GAGCGAAGACTTCTTCTACATCAAGGTCAGCGCAAGTCCAGCTCCAAAATGGGCAA 168  
 DB 912 GGCCATCGCTTTCGTGACCAACCAAGCAAGCGCGGCGGCACGGC 957



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:27:06 ; Search time 20.2431 Seconds  
(without alignments)

5226.648 Million cell updates/sec

Title: US-09-513-888C-1\_COPY\_112\_456

Perfect score: 345

Sequence: 1 tcgcagctacaagctgcgca.....gcaggagaagcaaatggg 345

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.6	11.5	1981	3	US-09-017-706-3
2	39.6	11.5	1981	3	US-09-017-706-4
3	39.6	11.5	1981	3	US-09-017-706-5
4	39.6	11.5	1981	3	US-09-017-706-6
5	39.6	11.5	1981	3	US-09-017-706-7
6	39.6	11.5	1981	3	US-09-017-706-8
7	36.4	10.6	2267	4	US-08-679-645-25
8	36.4	10.6	4800	3	US-08-941-445A-4
9	34.6	10.0	5267	3	US-08-976-255-2
10	34	9.9	13842	4	US-09-105-537-30
11	34	9.9	36778	4	US-09-105-537-5
12	34	9.9	38506	3	US-09-320-878-19
13	34	9.9	43280	2	US-08-804-227C-1
14	33	9.6	4403765	4	US-09-103-840A-2
15	33	9.6	4411529	4	US-09-103-840A-1
16	32.8	9.5	1590	4	US-09-504-358-21
17	32.8	9.5	1590	4	US-09-954-314-21
18	32.8	9.5	11471	4	US-09-504-358-16
19	32.8	9.5	11471	4	US-09-954-314-16
20	32.4	9.4	1293	2	US-08-924-440-1
21	32	9.3	3624	2	US-07-951-715A-6
22	32	9.3	3624	2	US-08-459-448A-6
23	32	9.3	3624	3	US-08-459-595A-6
24	32	9.3	3624	3	US-08-459-504B-6
25	32	9.3	3624	3	US-08-459-444-6
26	32	9.3	3624	3	US-09-053-549-7
27	32	9.3	3624	4	US-09-547-422-6

28	32	9.3	8854	3	US-09-053-549-1	Sequence 1, Appli	
29	31.8	9.2	768	4	US-09-235-451-6	Sequence 6, Appli	
30	31.8	9.2	3314	1	US-07-973-324A-5	Sequence 5, Appli	
31	31.8	9.2	3314	1	US-08-343-380-5	Sequence 5, Appli	
32	31.8	9.2	3314	4	US-09-072-435-5	Sequence 5, Appli	
33	31.8	9.2	3314	4	US-09-072-917A-5	Sequence 5, Appli	
C	34	31.6	9.2	1605	3	US-08-701-582D-1	Sequence 1, Appli
35	31.6	9.2	2086	1	US-07-973-324A-1	Sequence 1, Appli	
36	31.6	9.2	2086	1	US-08-343-380-1	Sequence 1, Appli	
37	31.6	9.2	2086	4	US-09-072-435-1	Sequence 1, Appli	
38	31.6	9.2	2086	4	US-09-072-917A-1	Sequence 1, Appli	
39	31.4	9.1	999	2	US-08-761-277A-46	Sequence 46, Appli	
40	31.4	9.1	4403765	4	US-09-103-840A-2	Sequence 2, Appli	
41	31.4	9.1	4411529	4	US-09-103-840A-1	Sequence 1, Appli	
42	31.2	9.0	606	2	US-08-432-871C-48	Sequence 48, Appli	
43	31.2	9.0	606	4	US-09-270-956-48	Sequence 48, Appli	
C	44	31	9.0	3366	4	US-09-171-461-27	Sequence 27, Appli
45	31	9.0	43804	4	US-09-171-461-1	Sequence 1, Appli	

ALIGNMENTS

RESULT 1  
US-09-017-706-3  
; Sequence 3, Application US/09017706A  
; Patent No. 6087147  
; GENERAL INFORMATION:  
; APPLICANT: ITO, YOSHIFUMI  
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING  
; TITLE OF INVENTION: MALTOPEPTAASE, VECTOR CONTAINING SAID GENE AND  
; FILE REFERENCE: 8361-0003-0  
; CURRENT APPLICATION NUMBER: US/09/017,706A  
; CURRENT FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: JP 305071/1997  
; EARLIER FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1981  
; TYPE: DNA  
; ORGANISM: Pseudomonas sp., Strain KO-8940  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1848)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (85)..(1848)  
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E  
; FEATURE:  
; OTHER INFORMATION: MUTATED GENOMIC DNA  
; FEATURE:  
; OTHER INFORMATION: PLASMID pos3410F57  
US-09-017-706-3

Query Match	11.5%	Score 39.6;	DB 3;	Length 1981;
Best Local Similarity	52.4%	Pred. No. 0.017;	79;	Indels 0;
Matches	87;	Conservative	0;	Mismatches 0;
Gaps	0;			
Qy	3	CGAGTACAGCTGGCGGAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGGCT	62	
Db	792	GCAGGTACAGCTGGCGGAGTCTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGGCT	851	
Qy	63	GCTGAGGTTGGCTTCTCCAGAGTCTCGGTTCAGGCAAGTCCAGCTCCAAAATGGGCA	122	
Db	852	CGGCAAGTGGCGGAGTCTCGGTTCAGGCAAGTCCAGCTCCAAAATGGGCA	911	
Qy	123	GAGCGAGACTTCTTCTACATCAAGGTCAGCCAGCAAGCCCGGGGC	168	
Db	912	GGCATCGTCTTCTGTCGACCAACGACGCGGCGGCCCGGC	957	

## RESULT 2

US-09-017-706-4  
; Sequence 4, Application US/09017706A  
; Patent No. 6087147  
; GENERAL INFORMATION:  
; APPLICANT: ITO, YOSHIFUMI  
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING  
; TITLE OF INVENTION: MALTOSE, VECTOR CONTAINING SAID GENE AND  
; TITLE OF INVENTION: TRANSFORMANT  
; FILE REFERENCE: 8361-0003-0  
; CURRENT APPLICATION NUMBER: US/09/017,706A  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: JP 305071/1997  
; EARLIER FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1981  
; TYPE: DNA  
; ORGANISM: Pseudomonas sp., Strain KO-8940  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1848)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (85)..(1848)  
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E  
; OTHER INFORMATION: MUTATED GENOMIC DNA  
; OTHER INFORMATION: PLASMID POS3410H57  
US-09-017-706-4

Query Match 11.5%; Score 39.6; DB 3; Length 1981;  
Best Local Similarity 52.4%; Pred. No. 0.017;  
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
  
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62  
DB 792 GCAGGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 851  
  
QY 63 GCTGAGGTTTGGCTTCTCCAGAGCTCCCGTCAAGAGCTCAACCGGTATTCGACGG 122  
DB 852 CGGCAAGCTGGCGACCTCGACCTTCGCGCCAGCTGGAACCTGATGCCACGAGCAA 911  
  
QY 123 GAGCGAAGACTTCTTCTACATCAAGTCAAGAGCTCAACCGGTATTCGACGGGCT 168  
DB 912 GGCCATCGCTTTCGTGACAACTCAAGAGCTCAACCGGTATTCGACGGGCT 957

## RESULT 3

US-09-017-706-5  
; Sequence 5, Application US/09017706A  
; Patent No. 6087147  
; GENERAL INFORMATION:  
; APPLICANT: ITO, YOSHIFUMI  
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING  
; TITLE OF INVENTION: MALTOSE, VECTOR CONTAINING SAID GENE AND  
; TITLE OF INVENTION: TRANSFORMANT  
; FILE REFERENCE: 8361-0003-0  
; CURRENT APPLICATION NUMBER: US/09/017,706A  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: JP 305071/1997  
; EARLIER FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1981  
; TYPE: DNA  
; ORGANISM: Pseudomonas sp., Strain KO-8940  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1848)

## ; FEATURE:

; NAME/KEY: mat\_peptide  
; LOCATION: (85)..(1848)  
; OTHER INFORMATION: METHOD OF DETERMINING CHARACTERISTIC: E  
; FEATURE:  
; OTHER INFORMATION: MUTATED GENOMIC DNA  
; FEATURE:  
; OTHER INFORMATION: PLASMID POS3410OL57  
US-09-017-706-5

Query Match 11.5%; Score 39.6; DB 3; Length 1981;  
Best Local Similarity 52.4%; Pred. No. 0.017;  
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
  
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62  
DB 792 GCAGGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 851  
  
QY 63 GCTGAGGTTTGGCTTCTCCAGAGCTCCCGTCAAGAGCTCAACCGGTATTCGACGG 122  
DB 852 CGGCAAGCTGGCGACCTCGACCTTCGCGCCAGCTGGAACCTGATGCCACGAGCAA 911  
  
QY 123 GAGCGAAGACTTCTTCTACATCAAGTCAAGAGCTCAACCGGTATTCGACGGGCT 168  
DB 912 GGCCATCGCTTTCGTGACAACTCAAGAGCTCAACCGGTATTCGACGGGCT 957

## RESULT 4

US-09-017-706-6  
; Sequence 6, Application US/09017706A  
; Patent No. 6087147  
; GENERAL INFORMATION:  
; APPLICANT: ITO, YOSHIFUMI  
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING  
; TITLE OF INVENTION: MALTOSE, VECTOR CONTAINING SAID GENE AND  
; TITLE OF INVENTION: TRANSFORMANT  
; FILE REFERENCE: 8361-0003-0  
; CURRENT APPLICATION NUMBER: US/09/017,706A  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: JP 305071/1997  
; EARLIER FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1981  
; TYPE: DNA  
; ORGANISM: Pseudomonas sp., Strain KO-8940  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1848)  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (85)..(1848)  
; FEATURE:  
; OTHER INFORMATION: MUTATED GENOMIC DNA  
; FEATURE:  
; OTHER INFORMATION: PLASMID: POS3410F139  
US-09-017-706-6

Query Match 11.5%; Score 39.6; DB 3; Length 1981;  
Best Local Similarity 52.4%; Pred. No. 0.017;  
Matches 87; Conservative 0; Mismatches 79; Indels 0; Gaps 0;  
  
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62  
DB 792 GCAGGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 851  
  
QY 63 GCTGAGGTTTGGCTTCTCCAGAGCTCCCGTCAAGAGCTCAACCGGTATTCGACGG 122  
DB 852 CGGCAAGCTGGCGACCTCGACCTTCGCGCCAGCTGGAACCTGATGCCACGAGCAA 911  
  
QY 123 GAGCGAAGACTTCTTCTACATCAAGTCAAGAGCTCAACCGGTATTCGACGGGCT 168



```

RESULT 8
US-08-941-445A-4
; Sequence 4, Application US/08941445A
; Patent No. 6107060
; GENERAL INFORMATION:
; APPLICANT: Keeling, Peter
; APPLICANT: Guan, Hanping
; TITLE OF INVENTION: Starch Encapsulation
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle
; CITY: Boulder
; STATE: CO
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/941,445A
; FILING DATE: 30-SEP-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,855
; FILING DATE: 30-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Winner, Ellen P
; REGISTRATION NUMBER: 28,547
; REFERENCE/DOCKET NUMBER: 89-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 10.0%; Score 34.6; DB 3; Length 5267;  
Best Local Similarity 52.4%; Pred. No. 0.69;  
Matches 76; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 58 GGGCTGCTGAGTTGGCTTCTCCAGGACTCCGGTCAAGGCAAGTCCAGCTCCAAATG 117  
DB 2958 GAGGTGGAGGCGGTGTCTCGGGCCGGGGCCCTACCTCTGAGCGAGCTCCGCAAG 2899

QY 118 GCAAGAGCGAAGACTTCTTACATCAAGGTGAGCAGAAAGCCGGGGCTCCCATCAC 177  
DB 2898 GCCTGGGGCTACACCCCTTCTTGGCGCTCTCTTGGAGCAGAACTCAGGGGACTCATAGTTC 2839

QY 178 CCAGATTACAGCGACTGTCCAGCG 202  
DB 2838 TCGGTGTATAGCCACTGTCCAGCG 2814

RESULT 10  
US-09-105-537-30  
; Sequence 30, Application US/09105537A  
; Patent No. 6265202  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/105,537A  
; CURRENT FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 30  
; LENGTH: 13842  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-105-537-30

Query Match 9.9%; Score 34; DB 4; Length 13842;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGACTCCGGTCAAGGCAAGTCCAGTCCAAATGGGCAAGAGCAAGACTTCTTCT 139  
DB 5951 CCGAGCGCTACTGGCTCGAGAACAACCTCCCGCGCCCTGCGCACCGGCGACGACTGGCGCT 6010

QY 140 ACATCAAGTTCAGCCAGAAAGCCCGGGCTCCCATCCAGATTACAGGCACTGTCCA 199  
DB 6011 ACCGATCGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGGCGCTGTCGG 6070

QY 200 GCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTTGACCCGTCACACCC 249  
DB 6071 GCGGCTGGCTCGCGCTACGCGCGGAGGACCACTCCGCGGAGCGCGCGCCGCC 6120

RESULT 11  
US-09-105-537-5  
; Sequence 5, Application US/09105537A  
; Patent No. 6265202  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, D.H.  
; APPLICANT: Liu, H.  
; APPLICANT: Xue, Y.  
; APPLICANT: Zhao, L.  
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
; FILE REFERENCE: 600.438US1  
; CURRENT APPLICATION NUMBER: US/09/105,537A  
; CURRENT FILING DATE: 1998-06-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 36778  
; TYPE: DNA

; ORGANISM: Streptomyces venezuelae  
US-09-105-537-5

Query Match 9.9%; Score 34; DB 4; Length 36778;  
Best Local Similarity 50.0%; Pred. No. 2;  
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGACTCCGGTCAAGGCAAGTCCAGTCCAAATGGGCAAGAGCAAGACTTCTTCT 139  
DB 7692 CCGAGCGCTACTGGCTCGAGAACAACCTCCCGCGCCCTGCGCACCGGCGACGACTGGCGCT 7751

QY 140 ACATCAAGTTCAGCCAGAAAGCCCGGGCTCCCATCCAGATTACAGGCACTGTCCA 199  
DB 7752 ACCGATCGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGGCGCTGTCGG 7811

QY 200 GCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTTGACCCGTCACACCC 249  
DB 7812 GCGGCTGGCTCGCGTCAAGCGGAGGACCACTCCGCGAGCGCGCGCC 7861

RESULT 12  
US-09-320-878-19  
; Sequence 19, Application US/09320878A  
; Patent No. 6117659  
; GENERAL INFORMATION:  
; APPLICANT: ASHLEY, Gary  
; APPLICANT: BETLACH, Melanie C.  
; APPLICANT: BETLACH, Mary C.  
; APPLICANT: MCDANIEL, Robert  
; APPLICANT: TANG, Li  
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE  
; FILE REFERENCE: 300622002120  
; CURRENT APPLICATION NUMBER: US/09/320,878A  
; CURRENT FILING DATE: 1999-05-27  
; EARLIER APPLICATION NUMBER: CIP OF 09/141,908  
; EARLIER FILING DATE: 1998-08-28  
; EARLIER APPLICATION NUMBER: CIP OF 09/073,538  
; EARLIER FILING DATE: 1998-05-06  
; EARLIER APPLICATION NUMBER: CIP OF 08/846,247  
; EARLIER FILING DATE: 1997-04-30  
; EARLIER APPLICATION NUMBER: 60/119,139  
; EARLIER FILING DATE: 1999-02-08  
; EARLIER APPLICATION NUMBER: 60/100,880  
; EARLIER FILING DATE: 1998-09-22  
; EARLIER APPLICATION NUMBER: 60/087,080  
; EARLIER FILING DATE: 1998-05-28  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 19  
; LENGTH: 38506  
; TYPE: DNA  
; ORGANISM: Streptomyces venezuelae  
US-09-320-878-19

Query Match 9.9%; Score 34; DB 3; Length 38506;  
Best Local Similarity 50.0%; Pred. No. 2.1;  
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGACTCCGGTCAAGGCAAGTCCAGTCCAAATGGGCAAGAGCAAGACTTCTTCT 139  
DB 5834 CCGAGCGCTACTGGCTCGAGAACAACCTCCCGCGCCCTGCGCACCGGCGACGACTGGCGCT 5893

QY 140 ACATCAAGTTCAGCCAGAAAGCCCGGGCTCCCATCCAGATTACAGGCACTGTCCA 199  
DB 5894 ACCGATCGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGGCGCTGTCGG 5953

QY 200 GCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTTGACCCGTCACACCC 249  
DB 5954 GCGGCTGGCTCGCGTCAAGCGCGGAGGACCACTCCGCGGAGCGCGCGCCGCC 6003

RESULT 13  
US-08-804-227C-1

; Sequence 1, Application US/08804227C  
; Patent No. 5876991  
; GENERAL INFORMATION:  
; APPLICANT: Dehoff, Bradley S.  
; APPLICANT: Kuhstoss, Stuart A.  
; APPLICANT: Rostock, Paul R., Jr.  
; APPLICANT: Sutton, Kimberly L.  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THOMAS G. PLANT 1501  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII(DOS) Text only  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804,227C  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas, G.  
; REGISTRATION NUMBER: 35,784  
; REFERENCE/DOCKET NUMBER: X-8231  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-2459  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43280 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 816..14234  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 14351..19945  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 20010..31199  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 31232..36067  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 36249..41774  
; US-08-804-227C-1

Query Match 9.9%; Score 34; DB 2; Length 43280;  
Best Local Similarity 59.2%; Pred. No. 2.1;  
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
Qy 213 GGGCCAGGCTGGGGTGGACTTTGACCGGTCCACACCCCAAGCTCATGCCCTTCTCCAA 272  
Db 15004 GGCACCTGGCGGGGGGACCGTGTCTCCACCCCAAGCTGTCTCGTGGACTTCTCCG 15063  
Qy 273 TCAGCTAGAAATGGTAAGCGGGGGTGGTGGCAAGGT 310  
Db 15064 CCAGCGCGTCTGGCCCCGACGCGCTGCAAGGCT 15101

RESULT 14  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:

; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
; US-09-103-840A-2

Query Match 9.6%; Score 33; DB 4; Length 4403765;  
Best Local Similarity 49.7%; Pred. No. 18;  
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
Qy 75 CTTCTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGCAAGCGAAGACTT 134  
Db 3167807 CCTCAACGAGGTCTATCTGCCATGGAATCCCGACTCGACGGTGTATCCGACGGGACAT 3167748  
Qy 135 CTTCTACATCAAGGTACAGCCAGCAAGAGCCGGGGCTCCCATCACCAGATTACACGGCACT 194  
Db 3167747 CGTCAACATCGACGTACCGCTACATCGTGGGGTGCACGGTGCACCAACGCGACGT 3167688  
Qy 195 GTCCAGCGGGGATTAGGGGCCAGGCTGGGGTGGACTTTGACCGTCC 243  
Db 3167687 TCCGGCGCGGATGTCCGACAGCAACACCGCTTCTGCTGACCGGACC 3167639

RESULT 15  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37Rv  
; US-09-103-840A-1

Query Match 9.6%; Score 33; DB 4; Length 4411529;  
Best Local Similarity 49.7%; Pred. No. 18;  
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
Qy 75 CTTCTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGCAAGCGAAGACTT 134  
Db 3173694 CCTCAACGAGGTCTATCTGCCATGGAATCCCGACTCGACGGTGTATCCGACGGGACAT 3173635  
Qy 135 CTTCTACATCAAGGTACAGCCAGAAAGCCGGGGTCCCATCACCAGATTACACGGCACT 194  
Db 3173634 CGTCAACATCGACGTACCGCTCATCTCGTGGGGTGCACGGTGCACCAACGCGACGT 3173575



Qy 195 GTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCGTCC 243  
Db 3173574 TCCGCCGGCGATGTCGAGACGAAACACCGGTTGCTCGTTGACCGGACC 3173526

Search completed: June 15, 2003, 03:45:44  
Job time : 36.2431 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 02:18:56 ; Search time 50.2221 Seconds  
(without alignments)  
9947.427 Million cell updates/sec

Title: US-09-513-888c-1\_COPY\_112\_456  
Perfect score: 345  
Sequence: 1 tcgcagtagaagctgcgca.....gcaggagaaagcaaatggg 345

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1029858 seqs, 72403093 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	33.3	2411	9	US-09-764-904-89
2	115	33.3	2411	9	US-10-091-548-89
3	115	33.3	2411	9	US-10-074-095-1133
4	115	33.3	2411	10	US-09-764-860-1133
5	36.4	10.6	2267	9	US-09-961-077-25
6	36	10.4	560	9	US-09-764-904-17
7	36	10.4	560	9	US-10-091-548-17
8	36	10.4	560	9	US-10-074-095-269
9	36	10.4	560	10	US-09-764-860-269
10	35	10.1	328	10	US-09-878-574-2638
11	34	9.9	13842	9	US-09-860-846-30
12	34	9.9	13842	9	US-09-988-384B-30
13	34	9.9	13842	9	US-09-836-821-30
14	34	9.9	13842	10	US-09-861-289-30
15	34	9.9	36778	9	US-09-860-846-5
16	34	9.9	36778	9	US-09-836-821-5
17	34	9.9	36778	10	US-09-861-289-5
18	34	9.9	37948	9	US-09-988-384B-5
19	34	9.9	38506	9	US-09-793-708-19

C	20	34	9.9	48841	9	US-09-844-653-32	Sequence 32, Appl
	21	33.2	9.6	2521	10	US-09-925-300-521	Sequence 521, Appl
	22	32.8	9.5	1590	10	US-09-954-314-21	Sequence 21, Appl
	23	32.8	9.5	11471	10	US-09-954-314-16	Sequence 16, Appl
	24	32.4	9.4	1883	12	US-10-033-109-9	Sequence 9, Appl
	25	32.4	9.4	2313	9	US-09-887-527-6	Sequence 6, Appl
	26	32.4	9.4	2412	9	US-10-037-270-881	Sequence 881, Appl
C	27	32.2	9.3	1985	10	US-09-864-761-15510	Sequence 15510, A
	28	32	9.3	2079	9	US-10-076-816-3	Sequence 3, Appl
	29	32	9.3	2079	9	US-10-077-381-3	Sequence 3, Appl
	30	32	9.3	3624	9	US-09-988-462-6	Sequence 6, Appl
	31	31.8	9.2	768	9	US-09-978-303-6	Sequence 6, Appl
C	32	31.8	9.2	926	9	US-10-198-846-13362	Sequence 13362, A
	33	31.4	9.1	999	9	US-09-925-664-46	Sequence 46, Appl
	34	31.4	9.1	1278	9	US-10-016-647-1	Sequence 1, Appl
	35	31.4	9.1	1844	9	US-10-016-647-3	Sequence 3, Appl
	36	31.4	9.1	3526	9	US-10-037-270-182	Sequence 182, App
	37	31.2	9.0	455	9	US-09-918-995-2482	Sequence 2482, Ap
	38	31.2	9.0	839	9	US-10-029-905-11	Sequence 11, Appl
	39	31.2	9.0	839	10	US-09-880-107-2223	Sequence 2223, Ap
	40	31	9.0	1380	10	US-09-925-301-318	Sequence 318, App
	41	31	9.0	1656	12	US-10-033-109-13	Sequence 13, Appl
	42	31	9.0	1919	9	US-10-168-066-17	Sequence 17, Appl
	43	31	9.0	2367	10	US-09-995-587A-2	Sequence 2, Appl
	44	31	9.0	2394	10	US-09-995-587A-3	Sequence 3, Appl
	45	31	9.0	2592	10	US-09-995-587A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1  
US-09-764-904-89  
; Sequence 89, Application US/09764904  
; Patent No. US20020173454A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA122  
; CURRENT APPLICATION NUMBER: US/09/764,904  
; PRIORITY FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 89  
; LENGTH: 2411  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-904-89

Query Match 33.3%; Score 115; DB 9; Length 2411;  
Best Local Similarity 100.0%; Pred. No. 9e-28;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	231	CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTCAATCAGCTAGAAATGGTAAG	290
Db	1	CTTTGACCCGTCACACCCCAAGCTCATGCCCTTCTCAATCAGCTAGAAATGGTAAG	60
QY	291	CGGGGTCGCTGGCAGGTAAGTGGTGGAAACCGAGGAAAGCAAAATGGG	345
Db	61	CGGGGTCGCTGGCAGGTAAGTGGTGGAAACCGAGGAAAGCAAAATGGG	115

RESULT 2  
US-10-091-548-89  
; Sequence 89, Application US/10091548  
; Publication No. US20030049703A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA122C1  
; CURRENT APPLICATION NUMBER: US/10/091,548  
; PRIORITY FILING DATE: 2002-03-07

```

; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-89

Query Match      33.3%; Score 115; DB 9; Length 2411;
Best Local Similarity 100.0%; Pred. No. 9e-28;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      231 CTTTGACCGTCCACACCCCAAGCTCATGCCCTTCCCAATCAGCTAGAAATGGTAA 290
Db      1 CTTTGACCGTCCACACCCCAAGCTCATGCCCTTCCCAATCAGCTAGAAATGGTAA 60

QY      291 CGGGGGTGGCTGGCAAGGTAAAGTGGGTTGGAACCGCAGGAGAAAGCAAAATGGG 345
Db      61 CGGGGGTGGCTGGCAAGGTAAAGTGGGTTGGAACCGCAGGAGAAAGCAAAATGGG 115

RESULT 3
US-10-074-095-1133
; Sequence 1133, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC080C1
; CURRENT APPLICATION NUMBER: US/10/074,095
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: 09/764,860
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17

```

;; PRIOR APPLICATION NUMBER: 60/226,681  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,759  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/225,213  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/227,182  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,214  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/225,836  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: 60/230,438  
;; PRIOR FILING DATE: 2000-09-06  
;; PRIOR APPLICATION NUMBER: 60/215,135  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: 60/225,266  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/249,218  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,208  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,213  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,212  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,207  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,245  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,244  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,217  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,211  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,215  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,264  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,214  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,297  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/232,400  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/231,242  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,081  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,080  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,414  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,244  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/233,064  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/233,063  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,397  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,399  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,401  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/241,808  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,826  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,786  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,221

;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,475  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/231,243  
;; PRIOR FILING DATE: 2000-09-08  
  
Query Match 33.3%; Score 115; DB 9; Length 2411;  
Best Local Similarity 100.0%; Pred. No. 9e-28;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 231 CTTTGACCCGTCACACACCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAGAAATGGTAAG 290  
|||||  
DB 1 CTTTGACCCGTCACACACCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAGAAATGGTAAG 60  
|||||  
  
QY 291 CGGGGGTCGCTGGCAAGGTAAGTGGTTGGAACGCGAGGAGAAAGCAAAATGGG 345  
|||||  
DB 61 CGGGGGTCGCTGGCAAGGTAAGTGGTTGGAACGCGAGGAGAAAGCAAAATGGG 115  
|||||  
  
RESULT 4  
US-09-764-860-1133  
; Sequence 1133, Application US/09764860  
; Patent No. US20020094953A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008  
; CURRENT APPLICATION NUMBER: US/09/764,860  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1198  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1133  
; LENGTH: 2411  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-860-1133  
  
Query Match 33.3%; Score 115; DB 10; Length 2411;  
Best Local Similarity 100.0%; Pred. No. 9e-28;  
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 231 CTTTGACCCGTCACACACCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAGAAATGGTAAG 290  
|||||  
DB 1 CTTTGACCCGTCACACACCCCAAGCTCATGCCCTTCTCCAAATCAGCTAGAGAAATGGTAAG 60  
|||||  
  
QY 291 CGGGGGTCGCTGGCAAGGTAAGTGGTTGGAACGCGAGGAGAAAGCAAAATGGG 345  
|||||  
DB 61 CGGGGGTCGCTGGCAAGGTAAGTGGTTGGAACGCGAGGAGAAAGCAAAATGGG 115  
|||||  
  
RESULT 5  
US-09-961-077-25  
; Sequence 25, Application US/09961077  
; Publication No. US20030014775A1  
; GENERAL INFORMATION:  
; APPLICANT: Zwick, Michael G.  
; Edington, Brent E.  
; McSwiggen, James A.  
; Merlo, Patricia Ann Owens  
; Guo, Lining  
; Skokut, Thomas A.  
; Young, Scott A.  
; Folkerts, Otto  
; Merlo, Donald J.  
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR  
; MODULATION OF GENE EXPRESSION  
; IN PLANTS  
; NUMBER OF SEQUENCES: 1263  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lyon & Lyon  
; STREET: 633 West Fifth Street  
; Suite 4700

CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/961,077  
FILING DATE: 21-Sep-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/679,645  
FILING DATE: July 12, 1996  
APPLICATION NUMBER: 60/001,135  
FILING DATE: July 13, 1995  
APPLICATION NUMBER: 08/300,726  
FILING DATE: September 2, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 219/247  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2267 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-961-077-25  
Query Match 10.6%; Score 36.4; DB 9; Length 2267;  
Best Local Similarity 51.2%; Pred. No. 0.057;  
Matches 85; Conservative 0; Mismatches 81; Indels 0; Gaps 0;  
QY 3 GCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGGCT 62  
DB 1094 GAAGCCGGATCTCGAGCGCAGACGGTCTCTACCGTCAGCCCTACTAGCCGAGA 1153  
QY 63 GGTGAGTTTGCTTCTCCAGGACTCCGGTCAACGGCAAGTCCAGCTCAAAATGGCAA 122  
DB 1154 GCTCATCTCCGGCATCGCCAGGGCTGCGAGCTCGACAAACATCATGCGCTCACCGCAT 1213  
QY 123 GAGCGAGACTTCTTACATCAAGTCAAGTCAGCCAGAAAGCCCGGGC 168  
DB 1214 CACCGGCATCGTCAACGGCATGAGCTCAGCGAGTGGGACCCGAC 1259  
RESULT 6  
US-09-764-904-17  
Sequence 17, Application US/09764904  
Patent No. US20020173454A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA122  
CURRENT APPLICATION NUMBER: US/09/764,904  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 17  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-904-17  
Query Match 10.4%; Score 36; DB 9; Length 560;  
Best Local Similarity 88.6%; Pred. No. 0.053;  
Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 251 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTAACGGG 294  
DB 1 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTCCGAG 44  
RESULT 7  
US-10-091-548-17  
Sequence 17, Application US/10091548  
Publication No. US20030049703A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA122C1  
CURRENT APPLICATION NUMBER: US/10/091,548  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 137  
Prior Application removed - See File Wrapper or Palm  
SEQ ID NO 17  
LENGTH: 560  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-091-548-17  
Query Match 10.4%; Score 36; DB 9; Length 560;  
Best Local Similarity 88.6%; Pred. No. 0.053;  
Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 251 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTAACGGG 294  
DB 1 CCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTCCGAG 44  
RESULT 8  
US-10-074-095-269  
Sequence 269, Application US/10074095  
Publication No. US2003007704A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC008C1  
CURRENT APPLICATION NUMBER: US/10/074,095  
CURRENT FILING DATE: 2002-02-14  
Prior Application Number: 09/764,860  
Prior Filing Date: 2001-01-17  
Prior Application Number: 60/179,065  
Prior Filing Date: 2000-01-31  
Prior Application Number: 60/180,628  
Prior Filing Date: 2000-02-04  
Prior Application Number: 60/214,886  
Prior Filing Date: 2000-06-28  
Prior Application Number: 60/217,487  
Prior Filing Date: 2000-07-11  
Prior Application Number: 60/225,758  
Prior Filing Date: 2000-08-14  
Prior Application Number: 60/220,963  
Prior Filing Date: 2000-07-26  
Prior Application Number: 60/217,496  
Prior Filing Date: 2000-07-11  
Prior Application Number: 60/225,447  
Prior Filing Date: 2000-08-14  
Prior Application Number: 60/218,290  
Prior Filing Date: 2000-07-14  
Prior Application Number: 60/225,757  
Prior Filing Date: 2000-08-14  
Prior Application Number: 60/226,868  
Prior Filing Date: 2000-08-22  
Prior Application Number: 60/216,647

; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02

; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,080  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,414  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/231,244  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/233,064

; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 60/233,063  
 ; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 60/232,397  
 ; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 60/232,399  
 ; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 60/232,401  
 ; PRIOR FILING DATE: 2000-09-14  
 ; PRIOR APPLICATION NUMBER: 60/241,808  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/241,826  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/241,786  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/241,221  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/246,475  
 ; PRIOR FILING DATE: 2000-11-08  
 ; PRIOR APPLICATION NUMBER: 60/231,243  
 ; PRIOR FILING DATE: 2000-09-08

Query Match 10.4%; Score 36; DB 9; Length 560;  
 Best Local Similarity 88.6%; Pred. No. 0.053;  
 Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 251 CCAAGCTATGCGCTTCTCCCAATCAGCTAGAGAAATGGTAAGCGGG 294  
 DB 1 CCAAGCTATGCGCTTCTCCCAATCAGCTAGAGAAATGGTCCGAG 44

RESULT 9

US-09-764-860-269  
 ; Sequence 269, Application US/09764860  
 ; Patent No. US20020094953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PC008  
 ; CURRENT APPLICATION NUMBER: US/09/764,860  
 ; PRIOR FILING DATE: 2001-01-17  
 ; Prior application data removed - consult PALM or file wrapper  
 ; NUMBER OF SEQ ID NOS: 1198  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 269  
 ; LENGTH: 560  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-764-860-269

Query Match 10.4%; Score 36; DB 10; Length 560;  
 Best Local Similarity 88.6%; Pred. No. 0.053;  
 Matches 39; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 251 CCAAGCTATGCGCTTCTCCCAATCAGCTAGAGAAATGGTAAGCGGG 294  
 DB 1 CCAAGCTATGCGCTTCTCCCAATCAGCTAGAGAAATGGTCCGAG 44

RESULT 10

US-09-878-574-2638  
 ; Sequence 2638, Application US/09878574  
 ; Patent No. US20020110548A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Byrum, Joseph R.  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Thompson, Michael D.  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 ; FILE REFERENCE: 38-21(15401)B  
 ; CURRENT APPLICATION NUMBER: US/09/878,574  
 ; PRIOR FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: 09/333,535

; PRIOR FILING DATE: 1999-06-14  
 ; NUMBER OF SEQ ID NOS: 15775  
 ; SEQ ID NO 2638  
 ; LENGTH: 328  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; OTHER INFORMATION: Clone ID: LIB3028-022-Q1-B1-D5  
 US-09-878-574-2638

Query Match 10.1%; Score 35; DB 10; Length 328;  
 Best Local Similarity 54.2%; Pred. No. 0.098;  
 Matches 71; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 113 AAATGGCGCAGAGCAAGACTTCTTACATCAAGGTCAGCCAGAAAGCCCGGGGCTCCC 172  
 DB 161 AGAAGAGCAACAGCAGCAGACACTTTTGAGACAGGTAAGCCTCCAAATTCCTATCTCGC 220  
 QY 173 ATCACCAGGATTACACGGCAGCTGTCCAGCGGGGATTTAGGGGGCCAGGCTGGGGTGGACT 232  
 DB 221 TTGGAATGGAATCCAGAGCTCTGTGCGCATAGGCTGGCGGCTGGCGGTGATGGGCC 280  
 QY 233 TTGACCCCGTCC 243  
 DB 281 AAAACCTAGCC 291

RESULT 11

US-09-860-846-30  
 ; Sequence 30, Application US/09860846  
 ; Patent No. US20020164742A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sherman, D.H.  
 ; APPLICANT: Liu, H.  
 ; APPLICANT: Xue, Y.  
 ; APPLICANT: Zhao, L.  
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin  
 ; FILE REFERENCE: 600.438US1  
 ; CURRENT APPLICATION NUMBER: US/09/860,846  
 ; CURRENT FILING DATE: 2001-05-18  
 ; PRIOR APPLICATION NUMBER: 09/105,537  
 ; PRIOR FILING DATE: 1998-06-26  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 30  
 ; LENGTH: 13842  
 ; TYPE: DNA  
 ; ORGANISM: Streptomyces venezuelae  
 US-09-860-846-30

Query Match 9.9%; Score 34; DB 9; Length 13842;  
 Best Local Similarity 50.0%; Pred. No. 0.56;  
 Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 80 CCCAGGACTCCGGTCACGCAAGTCCAGCTCCAAAATGGCAAGAGCGAAGACTTCTTCT 139  
 DB 5951 CCGAGCGCTACTGGCTCGAGAACACTCCCGCGCGCCCTGSCCAGCGGCGACGACTGGCGT 6010  
 QY 140 ACATCAAGGTCAGCCAGAAAGCCCGGGGCTCCCATCCAGCATTTACAGGCACTGTCCA 199  
 DB 6011 ACCGCATCGACTGGAAGCGCCTCCCGGCGCGGAGGGTCCGAGCGCACCGGCTGTCCG 6070  
 QY 200 GCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTTGACCCGTCCACACCC 249  
 DB 6071 GCCGCTGCTCGCGTACGCGGAGGAGACACTCCGCGCAGCGCCGCCGCC 6120

RESULT 12

US-09-988-384B-30  
 ; Sequence 30, Application US/09988384B  
 ; Publication No. US20030073824A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sherman, D.H.  
 ; APPLICANT: Liu, H.



Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Indels	Gaps
US-09-861-289-30	9.9%;	50.0%;	DB 9;	13842;	85;	0;	0;
Sequence 30, Application US/09861289							
Patent No. US20020110897A1							
GENERAL INFORMATION:							
APPLICANT: Sherman, D.H.							
APPLICANT: Liu, H.							
APPLICANT: Xue, Y.							
APPLICANT: Zhao, L.							
TITLE OF INVENTION: DNA encoding methymycin and pikromycin							
FILE REFERENCE: 600.438US1							
CURRENT APPLICATION NUMBER: US/09/861,289							
CURRENT FILING DATE: 2001-05-18							
PRIOR APPLICATION NUMBER: 09/105,537							
PRIOR FILING DATE: 1998-06-26							
NUMBER OF SEQ ID NOS: 43							
SOFTWARE: FastSeq for Windows Version 3.0							
SEQ ID NO 30							
LENGTH: 13842							
TYPE: DNA							
ORGANISM: Streptomyces venezuelae							
US-09-861-289-30							
Query Match	9.9%;	Score 34;	DB 10;	Length 13842;			
Best Local Similarity	50.0%;	Pred. No. 0.56;					
Mismatches	85;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;		
QY 80	CCGAGGACTCCGGTCACGCAAGTCCAGCTCCAAATGGCAAGGCGAAGACTTCTTCT	139					
Db 5951	CCGAGGCTACTGGCTCGAGACACTCCGGCGCCCTGCCACCGGCGAGCACTGGCGCT	6010					
QY 140	ACATCAAGTACAGCAGAAAGCCCGGGCTCCCATCACCCAGATTACACGCACTGTCCA	199					
Db 6011	ACCGCATGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG	6070					
QY 200	CGGGGATTTAGGGGCCAGGCTGGGTGGACTTTGACCCGTCCACACCC	249					
Db 6071	GCCTGGCTCGCCCTCAGCGGAGGACCACTCCGCGAGCGCGCGCC	6120					
RESULT 13							
US-09-836-821-30							
Sequence 30, Application US/09836821							
Publication No. US20030087405A1							
GENERAL INFORMATION:							
APPLICANT: Sherman, D.H.							
APPLICANT: Liu, H.							
APPLICANT: Xue, Y.							
APPLICANT: Zhao, L.							
TITLE OF INVENTION: DNA encoding methymycin and pikromycin							
FILE REFERENCE: 600.438US1							
CURRENT APPLICATION NUMBER: US/09/836,821							
CURRENT FILING DATE: 2001-04-17							
PRIOR APPLICATION NUMBER: 09/105,537							
PRIOR FILING DATE: 1998-06-26							
NUMBER OF SEQ ID NOS: 43							
SOFTWARE: FastSeq for Windows Version 3.0							
SEQ ID NO 30							
LENGTH: 13842							
TYPE: DNA							
ORGANISM: Streptomyces venezuelae							
US-09-836-821-30							
Query Match	9.9%;	Score 34;	DB 9;	Length 13842;			
Best Local Similarity	50.0%;	Pred. No. 0.56;					
Mismatches	85;	Conservative 0;	Mismatches 85;	Indels 0;	Gaps 0;		
QY 80	CCGAGGACTCCGGTCACGCAAGTCCAGCTCCAAATGGCAAGGCGAAGACTTCTTCT	139					
Db 5951	CCGAGGCTACTGGCTCGAGACACTCCCGCGCCCTGCCACCGGCGAGCACTGGCGCT	6010					
QY 140	ACATCAAGTACAGCAGAAAGCCCGGGCTCCCATCACCCAGATTACACGCACTGTCCA	199					
Db 6011	ACCGCATGACTGGAAGCGCTCCCGCGCGCGAGGGTCCGAGCGCACCGCGCTGTCCG	6070					
QY 200	CGGGGATTTAGGGGCCAGGCTGGGTGGACTTTGACCCGTCCACACCC	249					
Db 6071	GCCTGGCTCGCCCTCAGCGGAGGACCACTCCGCGAGCGCGCGCC	6120	</				

Db 7752 ACCGATCGACTGGAAGCGCCTCCCGCGCGCCGAGGGTCCGAGCGCACCGGCTGTCCG 7811  
Qy 200 GCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTGACCCGTTCCACACCC 249  
Db 7812 GCGGCTGGCTCGCCGCTCAGCGGGAGGACCACTCCGCGCAGGCGCGCGCC 7861

Search completed: June 15, 2003, 06:49:10  
Job time : 52.2221 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 480.339 Seconds  
(without alignments)

11632.279 Million cell updates/sec

Title: US-09-513-888C-1\_COPY\_112\_456

Perfect score: 345

Sequence: 1 tcgcagtcacagctgcgcaa.....gcaggagaagcaaatggg 345

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

E8T.\*

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: gb\_gss.\*

18: em\_gss\_hum.\*

19: em\_gss\_inv.\*

20: em\_gss\_pln.\*

21: em\_gss\_vrt.\*

22: em\_gss\_fun.\*

23: em\_gss\_mam.\*

24: em\_gss\_mus.\*

25: em\_gss\_othr.\*

26: em\_gss\_pro.\*

27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	232.8	67.5	658	17 AZ858490	AZ858490 2M0163L10
C 2	232.4	67.4	472	12 BF706253	BF706253 280459 MA
C 3	220	63.8	596	10 AV663618	AV663618 AV663618
C 4	158.8	46.0	312	17 AZ595691	AZ595691 IM0408A04
C 5	122.2	35.4	817	9 A1984777	A1984777 wr85b07.x
C 6	99.2	28.8	668	12 BF058214	BF058214 7k28d05.x

7	62.6	18.1	72	10	BE048251
8	38.8	11.2	457	10	AV629798
9	38.4	11.1	366	13	BI064750
C 10	38	11.0	477	12	RG792715
C 11	37.6	10.9	759	14	BM675114
C 12	37	10.7	521	12	EG280616
C 13	37	10.7	535	12	EG280615
14	36.8	10.7	565	13	BJ268837
15	36.6	10.6	601	9	AL594733
16	36.6	10.6	624	9	AJ274159
17	36.6	10.6	644	9	AL783590
18	36.4	10.6	811	12	BF254153
19	36.4	10.6	2147	11	AY109531
20	35.8	10.4	540	9	AA240027
21	35.6	10.3	369	13	BJ316603
22	35.6	10.3	396	10	BE399579
23	35.6	10.3	433	9	AA485738
24	35.6	10.3	434	13	BJ213737
25	35.6	10.3	445	13	BJ222791
26	35.6	10.3	472	9	AL830730
27	35.6	10.3	476	10	BE400527
28	35.6	10.3	482	13	BM134430
29	35.6	10.3	519	9	AL830742
30	35.6	10.3	570	13	BJ289530
31	35.6	10.3	589	13	BM136270
32	35.6	10.3	600	13	BJ243353
33	35.6	10.3	604	12	EG604587
34	35.6	10.3	619	13	BJ209770
35	35.6	10.3	625	14	BQ609148
36	35.6	10.3	630	13	BJ289134
37	35.6	10.3	633	10	BE584071
38	35.6	10.3	645	14	BQ238903
39	35.6	10.3	649	13	BJ300787
40	35.6	10.3	658	13	BJ293180
41	35.6	10.3	669	13	BJ236056
42	35.6	10.3	728	13	BJ257903
43	35.6	10.3	814	9	AL820181
44	35.4	10.3	464	10	BE399162
C 45	35.4	10.3	476	13	BI340151

ALIGNMENTS

RESULT 1  
AZ858490/C  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AZ858490 658 bp DNA linear GSS 21-FEB-2001  
2M0163L10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC2M0163L10 R, DNA sequence.

AZ858490  
GI:13051689

GSS  
house mouse.

Mus musculus

1 (bases 1 to 658)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0163 row: L column: 10  
 Seq primer: CACACAGGAACAGTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 658.

## FEATURES

source

1. 658

/organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UUCG2M0163L10"  
 /clone\_lib="Mouse 10kb plasmid UUCG1M library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

140 a 173 c 177 g 168 t

BASE COUNT  
ORIGIN

Query Match 67.5%; Score 232.8; DB 17; Length 658;  
 Best Local Similarity 84.7%; Pred. No. 3.4e-57;

Matches 261; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1 TCGAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60  
 Db 522 TCACAGTACAAGCTCGGAAGTCTCCACCTCAAGAAGCTCAATCGATATCGATGG 463  
 Qy 61 CTGCTGAGTTGGCTTCTCCAGACTCCGCTACGCGCAAGTCCAGCTCCAAATGGC 120  
 Db 462 CTGCTGAGTTGGCTTCTCCAGACTCAGGCGGTGGAAGTCAAGTTCCAAATGGGA 403  
 Qy 121 AAGCGGAAGACTTCTTCTACATCAAGGTACGCCAAGAGCCGGGGCTCCCATCACCCA 180  
 Db 402 AAAAGCGAAGACTTCTTCTACATCAAGTACGCCAAGAGCCGGGGCTCCCATCGCCA 343  
 Qy 181 GATTACAGGCACTCTCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 240  
 Db 342 GACTACACAGCCCTCTCCAGTGGGACATAGGGGGTCAGACGGAGTAGATTTGATCCA 283  
 Qy 241 TCCACACCCCAAGCTATGCCCTTCTCCATCAGCTAGAAATGGTAAGGGGGTGGC 300  
 Db 282 GCCACCCACCAAGACTATGCCCTTCTCCATCACTGGAATGGTAGTAGGTTCTT 223  
 Qy 301 TGGCAAGG 308  
 Db 222 TGACAGGG 215

## RESULT 2

BF706253

LOCUS

BF706253 472 bp mRNA linear EST 25-APR-2001

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Bos taurus

## REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.  
 1 (bases 1 to 472)

Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Partea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.

TITLE

JOURNAL

MEDLINE

COMMENT

Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle  
 Genome Res. 11 (4), 626-630 (2001)  
 21180013  
 Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCCAGTCACGACG

Plate: 77 row: A column: 4

Seq primer: ATTTAGGTGACATATAG.

Location/Qualifiers

1. 472

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 3BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI; Library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle."

BASE COUNT 92 a 171 c 123 g 86 t

ORIGIN

Query Match 67.4%; Score 232.4; DB 12; Length 472;

Best Local Similarity 92.1%; Pred. No. 4e-57;

Matches 245; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 TCGAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 60

Db 207 TCGAGTACAGCTCGGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 266

Qy 61 CTGCTGAGTTGGCTTCTCCAGACTCCGCTACGCGCAAGTCCAGCTCCAAATGGC 120

Db 267 CTGCTGAGTTGGCTTCTCCAGACTCCGCGACATGCGAAGTCCAGTCCAAATGGC 326

Qy 121 AAGCGGAAGACTTCTTCTACATCAAGGTACGCCAAGAGCCGGGGCTCCCATCACCCA 180

Db 327 AAGCGGAAGACTTCTTCTACATCAAGGTACGCCAAGAGCCGGGGCTCCCATCGCCT 386

Qy 181 GATTACAGGCACTCTCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 240

Db 387 GATTACAGGCACTCTCAGTGGGGACCTAGGGGGCCAGGAGGAGTGGACTTTGACCCA 446

Qy 241 TCCACACCCCAAGCTCATGCCCTT 266

Db 447 TCCACACCCCAAGCTCATGCCCTT 472

RESULT 3

AV663618

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AV663618 596 bp mRNA linear EST 28-NOV-2001  
 AV663618 Bos taurus brain fetus Bos taurus cDNA clone E1BR037F07  
 5', mRNA sequence.  
 AV663618  
 AV663618.1 GI:9922648

KEYWORDS EUT. cow.  
SOURCE Bos taurus  
ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 596)  
AUTHORS Takasuga,A., Hirotsune,S., Itoh,R., Jithozono,A., Suzuki,H., Aso,H. and Sugimoto,Y.  
TITLE Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs

JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)  
MEDLINE 21570354  
COMMENT Contact: Yoshikazu Sugimoto  
Animal Genetics Division  
Shirakawa Institute of Animal Genetics  
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan  
Tel: 81-248-25-5641  
Fax: 81-248-25-5725  
Email: kazuugi@cocoa.ocn.ne.jp  
Single pass sequencing.  
This clone was obtained from a polyA-deleted cDNA library.

FEATURES  
source  
1..596  
Location/Qualifiers

/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone="E1BR037F07"  
/clone.lib="Bos taurus brain fetus"  
/tissue\_type="brain"  
/dev\_stage="fetus"  
/lab\_hosts="DH10B"  
/note="Vector: p2L1; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"  
BASE COUNT 114 a 196 c 170 g 103 t 13 others  
ORIGIN

Query Match 63.8%; Score 220; DB 10; Length 596;  
Best Local Similarity 91.0%; Pred. No. 1.8e-53;  
Matches 232; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 32 TCAGAGAGTCAACCGGTATTCCGACGGCTGCTGAGTTGGCTTCTCCAGAGCTCG 91  
Db 61 TCAAGAGCTCAACCGGTATTTCAGACGGCTGGTGGCTTCTCCAGAGCTCG 120

Qy 92 GTACGGCAAGTCCAGCTCCAAATGGCAAGAGCAAGACTTCTTACATCAAGGTCA 151  
Db 121 GACATGGCAAGTCCAGCTCCAAATGGCAAGAGCAAGACTTCTTACATCAAGGTCA 180

Qy 152 GCCAGAAAGCCGGGCTCCCATCACCCAGATTACCGGCACTGTCCAGCGGGATTAG 211  
Db 181 GCCAGAGGCCGGGGCTCCACCGCCCTGATTACCGGCACTGTCCAGTGGGNNNTAG 240

Qy 212 GGGGCCAGGCTGGGTGAGTTCACCGCTTCACACCCCGCCAGCTATGCCCTTCTCCA 271  
Db 241 GGGGCCAGGCGAGGTGAGTTCACCGCTTCACACCCCGCCAGCTATGCCCTTCTCCA 300

Qy 272 ATCAGCTAGAAATGG 286  
Db 301 ATCAGCTAGATGG 315

RESULT 4  
AZ595691/c  
LOCUS 312 bp DNA linear GSS 13-DEC-2000  
DEFINITION IM0408A04R Mouse 10kb plasmid UUC1M library Mus musculus genomic clone UUC1M0408A04 R, DNA sequence.

ACCESSION AZ595691  
VERSION AZ595691.1 GI:11717881  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 312)  
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0408 row: A column: 04  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 312.

FEATURES  
Location/Qualifiers  
1..312  
source

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUC1M0408A04"  
/clone.lib="Mouse 10kb plasmid UUC1M library"  
/sex="Male"  
/lab\_hosts="E. Coli strain XL10-Gold, Tl-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."  
BASE COUNT 65 a 71 c 101 g 75 t  
ORIGIN

Query Match 46.0%; Score 158.8; DB 17; Length 312;  
Best Local Similarity 88.7%; Pred. No. 1e-35;  
Matches 172; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 60  
Db 194 TCACAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAATCGATATCTAGATGG 135

Qy 61 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCAGTCCAAATGGGC 120  
Db 134 CTGCTGAGGTTTGGCTTCTCCAGGACTCAGGGCGTGGAAAGTCAAGTCCAAATGGGA 75

Qy 121 AAGCGGAAGACTTCTTACATCAAGGTGAGCGAGAAAGCCCGGGGCTCCATCACCCA 180  
Db 74 AAAAGCGAAGACTTCTTACATCAAGGTGAGCGAGAAAGCCCGGGGCTCCATCGCCA 15

Qy 181 GATTACACGGCACT 194  
Db 14 GACTACACGGCCT 1

```
RESULT 5
A1984777/c
LOCUS
DEFINITION
  A1984777.1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:2494453 3'
  similar to TR:O60299 O60299 KIAA0552 PROTEIN. ; mRNA sequence.
ACCESSION
  A1984777
VERSION
  A1984777.1 GI:5812054
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 817)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  DNA Sequencing by: Greg Lennon, Ph.D.
  NCI Library Arrayed by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www.bio.lnl.gov/bbrp/image/image.html
  Insert Length: 1307 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 428.
  Location/Qualifiers
  1..817
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:2494453"
  /clone_lib="NCI CGAP_Kid1"
  /lab_host="DH10B"
  /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
  a modified polylinker; Site 1: Not I; Site 2: Eco RI;
  Plasmid DNA from the normalized library NCI CGAP Kid3 was
  prepared, and ss circles were made in vitro. Following HAP
  purification, this DNA was used as tracer in a subtractive
  hybridization reaction. The driver was PCR-amplified cDNAs
  from a pool of 5,000 clones made from the same library
  (clones 132376-132391, 1456007-1456775, and
  1500552-1502855). Subtraction by Bento Soares and M.
  Fatima Bonaldo."
  133 a 235 c 266 g 180 t 3 others
BASE COUNT
ORIGIN
  Query Match 35.4%; Score 122.2; DB 9; Length 817;
  Best Local Similarity 86.4%; Pred. No. 7.3e-25;
  Matches 204; Conservative 0; Mismatches 24; Indels 8; Gaps 6;
  QY 52 TCCGACGGGCTGCTGAGTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCAGCTCC 111
  DB 756 TCCGACGGGCTGTAAGGT--GCTTTTCCGGACTCC-GTCACGGCAAGTCACTCCC 700
  QY 112 AAAATGGGCAAGAGGAGACTTCTTCTACATCAAGGTGAGCCAGAAAGCCGGGGCTCC 171
  DB 699 AAA--TGCAAGAGCGAGACTTCTTCTACATCAAGG-CAGCCAGAAAGCCGGGGCTCC 643
  QY 172 CATCACCGAGTATACAGGCATGTCCAGCGGGATTTAGGGGGCCAGGCT-GGGGTGGA 230
  DB 642 CATCACCGAGTATACAGG-CGACGGTCCAGCGGGATTTAGGGGGCCAGGTTGGGGGTGAA 584
  QY 231 CTTTGACCGTCCACACCCCAAGCTCATGCTTCTCCATCAGCTAGAAATGG 286
  DB 583 CTNTGCCCGTCCACCCCAAGGCTCATGCTTCTTCCATCAGCTAGAAATGG 528
  RESULT 6
A1984777/c
LOCUS
DEFINITION
  A1984777.1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:2494453 3'
  similar to TR:O60299 O60299 KIAA0552 PROTEIN. ; mRNA sequence.
ACCESSION
  A1984777
VERSION
  A1984777.1 GI:5812054
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 817)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps@mail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  DNA Sequencing by: Greg Lennon, Ph.D.
  NCI Library Arrayed by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www.bio.lnl.gov/bbrp/image/image.html
  Insert Length: 1307 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 428.
  Location/Qualifiers
  1..817
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:2494453"
  /clone_lib="NCI CGAP_Kid1"
  /lab_host="DH10B"
  /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
  a modified polylinker; Site 1: Not I; Site 2: Eco RI;
  Plasmid DNA from the normalized library NCI CGAP Kid3 was
  prepared, and ss circles were made in vitro. Following HAP
  purification, this DNA was used as tracer in a subtractive
  hybridization reaction. The driver was PCR-amplified cDNAs
  from a pool of 5,000 clones made from the same library
  (clones 132376-132391, 1456007-1456775, and
  1500552-1502855). Subtraction by Bento Soares and M.
  Fatima Bonaldo."
  133 a 235 c 266 g 180 t 3 others
BASE COUNT
ORIGIN
  Query Match 35.4%; Score 122.2; DB 9; Length 817;
  Best Local Similarity 86.4%; Pred. No. 7.3e-25;
  Matches 204; Conservative 0; Mismatches 24; Indels 8; Gaps 6;
  QY 52 TCCGACGGGCTGCTGAGTTGGCTTCTCCAGGACTCCGGTCACGGCAAGTCCAGCTCC 111
  DB 756 TCCGACGGGCTGTAAGGT--GCTTTTCCGGACTCC-GTCACGGCAAGTCACTCCC 700
  QY 112 AAAATGGGCAAGAGGAGACTTCTTCTACATCAAGGTGAGCCAGAAAGCCGGGGCTCC 171
  DB 699 AAA--TGCAAGAGCGAGACTTCTTCTACATCAAGG-CAGCCAGAAAGCCGGGGCTCC 643
  QY 172 CATCACCGAGTATACAGGCATGTCCAGCGGGATTTAGGGGGCCAGGCT-GGGGTGGA 230
  DB 642 CATCACCGAGTATACAGG-CGACGGTCCAGCGGGATTTAGGGGGCCAGGTTGGGGGTGAA 584
  QY 231 CTTTGACCGTCCACACCCCAAGCTCATGCTTCTCCATCAGCTAGAAATGG 286
  DB 583 CTNTGCCCGTCCACCCCAAGGCTCATGCTTCTTCCATCAGCTAGAAATGG 528
  RESULT 7
BE048251
LOCUS
DEFINITION
  BE048251.1 NCI CGAP Brn52 Homo sapiens cDNA clone IMAGE:2291902 5',
  mRNA sequence.
ACCESSION
  BE048251
```

```
BF058214/c
LOCUS
DEFINITION
  BF058214.1 NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:3476769 3'
  similar to TR:Q9YSW1 Q9YSW1 FEZ1. [5] TR:Q9YSW0 TR:Q9YSV9 TR:Q9YSV8
  TR:Q9YSV7 ; contains MER22.t3 TAR1 repetitive element ; mRNA
  sequence.
ACCESSION
  BF058214
VERSION
  BF058214.1 GI:10812110
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 668)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps@mail.nih.gov
  Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
  R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
  Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
  I.M.A.G.E. Consortium DNA Sequencing by: Washington University
  Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL, send email to:
  info@image.llnl.gov
  Seq primer: -40UP from Gibco
  High quality sequence stop: 441.
  Location/Qualifiers
  1..668
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone="IMAGE:3476769"
  /clone_lib="NCI CGAP Ov18"
  /tissue_type="fibroheoma"
  /lab_host="DH10B (phage-resistant)"
  /note="Organ: ovary; Vector: pT73D-Pac (Pharmacia) with a
  modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
  strand cDNA was primed with a Not I - oligo(dT) primer [5';
  TGTTCACATCTGAAGTGGAGCGCGCGCGACATTTTITTTTTTTT 3'];
  double-stranded cDNA was ligated to Eco RI adaptors
  (Pharmacia), digested with Not I and cloned into the Not
  I and Eco RI sites of the modified pT73 vector. Library
  went through one round of normalization, and was
  constructed by Bento Soares and M. Fatima Bonaldo."
  105 a 185 c 226 g 150 t 2 others
BASE COUNT
ORIGIN
  Query Match 28.8%; Score 99.2; DB 12; Length 668;
  Best Local Similarity 78.8%; Pred. No. 3.5e-18;
  Matches 130; Conservative 0; Mismatches 34; Indels 1; Gaps 1;
  QY 130 GACTTCTTCTACATCAAGGTGAGCCAGAAAGCCGGGGCTCCCATCCACCCAGATTACACG 189
  DB 668 GACTTCTTCTACATCATGTAGCCAGTAAGCCGAGGCTCCATCCCAAGATTATCG 609
  QY 190 GCACTGTCCAGCGGGGATTTAGGGGGCCAGGCTGGGTGGAGCTTTGACCCCTCCACACC 249
  DB 608 GCACGTGTCAGCGGAATAAGAGGGCCAGGATGGGTGAATATTGACCNCTCCACACCA 549
  QY 250 CCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGTAAGCGGG 294
  DB 548 CCCAAGCTCATGCCCTTCTCCATCAGCTAGAAATGTAAGCGGG 505
  RESULT 7
BE048251
LOCUS
DEFINITION
  BE048251.1 NCI CGAP Brn52 Homo sapiens cDNA clone IMAGE:2291902 5',
  mRNA sequence.
ACCESSION
  BE048251
```

DNA Res. 7 (5), 305-307 (2000)  
 20539644  
 Contact: Erika Asamizu  
 The First Laboratory for Plant Gene Research  
 Kazusa DNA Research Institute  
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan  
 Email: asamizuekazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.  
 Location/Qualifiers  
 1..457  
 /organism="Chlamydomonas reinhardtii"  
 /strain="C9"  
 /db\_xref="taxon:3055"  
 /clone="LC1066402.1"  
 /clone\_lib="Chlamydomonas reinhardtii 5% to 0.04% CO2"  
 /note="Vector: pBluescriptII SK; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from cells cultured  
 in a carbon stress acclimatized condition in which carbon  
 dioxide concentration in the bubbling gas was changed from  
 5% to 0.04%"

BASE COUNT      84 a    170 c    138 g    65 t

---

Query Match                  11.2%; Score 38.8; DB 10; Length 457;  
 Best Local Similarity       47.5%; Pred.No.1.3;  
 Matches 115; Conservative 0; Mismatched 127; Indels 0; Gaps 0;

```

QY      37  AAGCTCAACCGGTATTCCGACGGGTCTGCAGGTTTGGCTTTGCCAGGACTCCGGTCAC 96
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB     103  AAGGCCAAGGGGACAAGAAGTGATTTGGCCACGGGTTTATCGCGGCAGCCGGAG 162

QY      97  GGCAAGTCCAGTCTCCAAAATGGCGAAGACGAAGACTTCTTTACATCAAGGTTCAGCCAG 156
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB     163  GGCACCGCCACCACCCTGAAGCGCAACGGCTCGGACTTCTCGGCCACCATCATGGCGCG 222

QY     157  AAAGCCCGGGTCCCATTACCCAGNTTACAGGCACTGTCCAGCGGGGATTTAGGGGGC 216
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB     223  CTGTGCTGCTGGCGGCACATCACCATCTGGACGGAGCTGGAGCGGCGTGTACTCGGCGGAC 282

QY     217  CAGGCTGGGGTGGAGCTTTTGACCCGTCACACCCCCCAAGCTCATGGCCCTTCTCCAATCAG 276
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB     283  CCGCGCAAGGTGCCCCGAGCGCGTGTGCGCTGCCTCCATGACCTACCACGAGGCGCTGGAG 342

QY     277  CT 278
        ||
DB     343  CT 344
  
```

RESULT 9  
 BI064750/c  
 LOCUS  
 DEFINITION  
 protein [Homo sapiens] dbj|BAI74867.1| (AB020651) KIAA0844  
 accession  
 version  
 keywords  
 source  
 organism  
 chicken.  
 Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 366)  
 Cogburn,L.A., Morgan,R.W. and Burnside,J.  
 Chicken ESTs from fat  
 Unpublished (2001)  
 Contact: Larry A. Cogburn  
 University of Delaware  
 Townsend Hall, Newark, DE 19717, USA  
 Tel: 302-831-1335  
 Fax: 302-831-2822  
 Email: cogburn@udel.edu, www.chickest.udel.edu.  
 Location/Qualifiers

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

```

source
1. 366
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pgfin.pk002.g2"
/clone_lib="normalized chicken fat cDNA library"
/sex="Male and Female"
/tissue_type="fat"
/lab_host="E.coli EMPH10B"
/note="Vector: pSPORT1" 60 t

BASE COUNT      94 a   89 c   123 g   60 t
ORIGIN

Query Match      11.1%; Score 38.4; DB 13; Length 366;
Best Local Similarity 56.2%; Pred. No. 1.6; Mismatches 0; Indels 0; Gaps 0;
Matches 72; Conservative 0;

Qy 22 TCCTCCACCTCAAGAAGCTCAACGGTATTCCGACGGGCTGCTGAGGTTTGGCTTTCTCC 81
Db 283 TTCTCCAGATCCACAGCTGCAGCAGGTTCTCGATGAAGTCTGGAGTCTGGCTTTCTCC 224
Qy 82 CAGGACTCGGTCAGGCAAGTCCAGCTCCAAATGGCGAAGCAAGCAAGCTTTCTTAC 141
Db 223 CTTCACTTTCTTTCTGCTCTTCTTCGAGGAAGTGTGAAAGTAACAACCTTCTTCTC 164
Qy 142 ATCAAGGT 149
Db 163 CTTACGGT 156

RESULT 10
BG792715
LOCUS      BG792715          477 bp      mRNA      linear      EST 30-MAY-2001
DEFINITION UTSW H41A9-H21E10 UTSW Adult Mouse Cardiac Muscle Library Mus
            musculus cDNA clone UTSW_H41A9-H21E10, mRNA sequence.
ACCESSION  BG792715
VERSION     BG792715.1 GI:14128285
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 477)
AUTHORS     Gallardo,T.D., Schageman,J.V., Pertsemilidis,A., Garner,H.R.,
            Williams,R.S. and Shohet,R.V.
TITLE       UT Southwestern Medical Center, Adult Mouse Cardiac cDNA Library
JOURNAL     Unpublished (2001)
COMMENT     Contact: Schageman JJ
            Shohet/Garner Labs
            University of Texas Southwestern Medical Center
            6000 Harry Hines Blvd., NA2.226, Dallas, TX 75390, USA
            Tel: 214 648 1674
            Email: Jeff.Schageman@UTSouthwestern.edu
            cDNA library constructed by UTSW as a component of the Program for
            Genomic Applications (PGA) and the Reynolds Heart Disease
            Prevention grants for use in cDNA microarray experiments. Sequence
            Quality: Sequence ends were trimmed based on percentage of ambigu
            us base calls or 'N's in windowed segments. Sequencing: First-pass
            sequencing; ABI Prism 377 sequencer and analysis software.
            Seq primer: M13/pUC Reverse.
            Location/Qualifiers
            1. 477
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="UTSW H41A9-H21E10"
            /clone_lib="UTSW Adult Mouse Cardiac Muscle Library"
            /sex="Pooled"
            /tissue_type="Cardiac muscle"
            /dev_stage="2 months"
            /lab_host="DH5a"

source
1. 479
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahn-h-12-0-UI"
/clone_lib="UI-E-EJ0"
/tissue_type="fetal eyes, lens, eye anterior segment,
            optic nerve, retina, Retina Foveal and Macular, RPE and
            Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

Insertion site: TAGTCCACTGAATTCTGAGTG---. Other
information regarding entire library may be found at
http://pga.swmed.edu/Data/libraries/microarray_cdna_librar
ies.htm."

BASE COUNT      97 a   124 c   119 g   130 t   7 others
ORIGIN

Query Match      11.0%; Score 38; DB 12; Length 477;
Best Local Similarity 48.7%; Pred. No. 2.2; Mismatches 0; Indels 0; Gaps 0;
Matches 95; Conservative 0;

Qy 22 TCCTCCACCTCAAGAAGCTCAACGGTATTCCGACGGGCTGCTGAGGTTTGGCTTTCTCC 81
Db 221 TCCTGGCTCTCCAGAGGAGGAGCTGTAATTGAATCTGGTCTGGAGTATGTGCCGAG 280
Qy 82 CAGGACTCGGTCACGCAAGTCCAGCTCCAAATGGCGAAGCAAGCAAGCTTTCTTAC 141
Db 281 ACGGTGTACCGAGTGGCCGCCCACTTCACCAAGCCCAAGCTGATCACCCCTTTCATCTAC 340
Qy 142 ATCAAGCTCAGCCAGAAAGCCGGGCTCCCATCACCCAGATTACACGGCACTGTCCAGC 201
Db 341 ATCAAGGTANGCANGCANGCTGCTGGAGCACAACTGTCTGGGGTTCCTTCACTTCCCAAG 400
Qy 202 GGGGATTTAGGGGCG 216
Db 401 GTGGTCTTTGGGNC 415

RESULT 11
BG792715
LOCUS      BG792715          759 bp      mRNA      linear      EST 27-FEB-2002
DEFINITION UI-E-EJ0-ahn-h-12-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone
            UI-E-EJ0-ahn-h-12-0-UI 3', mRNA sequence.
ACCESSION  BG792715
VERSION     BG792715.1 GI:18985012
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 759)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Forward
            POLYA=Yes.
            Location/Qualifiers
            1. 759
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="UI-E-EJ0-ahn-h-12-0-UI"
            /clone_lib="UI-E-EJ0"
            /tissue_type="fetal eyes, lens, eye anterior segment,
            optic nerve, retina, Retina Foveal and Macular, RPE and
            Choroid"
            /dev_stage="fetal and adult"
            /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
            /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

```



modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGAATGACGA; eye anterior segment, AATGCCGCAI; optic nerve, CCAATTAAGTG; retina, CGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).

TAG LIB=UI-E-EJ0  
TAG\_TISSUE=human lens  
TAG\_SEQ=CGAATGACGA"

BASE COUNT 173 a 156 c 183 g 246 t 1 others  
ORIGIN

Query Match 10.9%; Score 37.6; DB 14; Length 759;  
Best Local Similarity 55.3%; Pred. No. 3.5;  
Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Qy 199 AGCGGGGATTAGGGGGCAGGCTGGGCTGGACTTTGACCCGTCACACACCCCAAGCTC 258  
|||||  
Db 704 AGCGACATCGGGACCCACACTCTGCTTCAGTTGCATCATCTCCACCCCAATTC 645  
|||||  
Qy 259 ATGCCCTTCTCAATCAGCTGAATGTGTAAGCGGGGTGCTGCGCAAGGGTAAAGTGGGT 318  
|||||  
Db 644 AACTCTGACCAATACAAAGATTTTAAACAGGATTTCTTTCGAGAAAGCTGACT 585  
|||||  
Qy 319 TGGAAACGAGG 330  
|||||  
Db 584 TGGAAACGCGG 573

RESULT 12  
BG280616  
LOCUS 521 bp mRNA linear EST 21-FEB-2001  
DEFINITION c5blinp.r1 Neurospora crassa sexual cDNA library, Uni-zap vector  
system Neurospora crassa cDNA clone c5blnp 5', mRNA sequence.  
ACCESSION BG280616  
VERSION BG280616.1 GI:13079199  
KEYWORDS EST.  
SOURCE Neurospora crassa.  
ORGANISM Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 521)  
AUTHORS Kupfer,D., Lai,H., Nelson,M. and Roe,B.  
TITLE ESTs from a Neurospora crassa Sexual cDNA Library  
JOURNAL Unpublished (2001)  
COMMENT Other ESTs: c5blinp.f1  
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
Department of Chemistry and Biochemistry  
Advanced Center for Genome Technology, University of Oklahoma  
620 Parrington Oval, Norman, OK 73019, USA  
Tel: 405 325 4912  
Fax: 405 325 7762  
Email: broe@ou.edu

FEATURES  
source  
1..521  
/organism="Neurospora crassa"  
/strain="wild type"  
/db\_xref="taxon:5141"

Query Match 10.9%; Score 37; DB 12; Length 535;  
Best Local Similarity 58.7%; Pred. No. 4.6;

BASE COUNT 118 a 130 c 143 g 144 t  
ORIGIN

/clone="c5blinp"  
/clone lib="Neurospora crassa sexual cDNA library, Uni-zap  
vector system"  
/tissue\_type="perithecia (fruiting bodies)"  
/dev\_stage="sexual"  
/lab\_host="E. coli strain SOLR"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript  
; 3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 130 a 154 c 115 g 122 t  
ORIGIN

Query Match 10.7%; Score 37; DB 12; Length 521;  
Best Local Similarity 58.7%; Pred. No. 4.6;  
Matches 64; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 30 CCTCAAGAGCTCAACGGTATTCGACGGGCTGCTGAGGTTCCTCCAGGACTC 89  
|||||  
Db 215 CATCATGTCGAGTAACAGGAATTACGATTTCTGTGCGCTTTAGCGAAGACTCATTC 274  
|||||  
Qy 90 CGGTCACGGCAAGTCCAGCTCCAAAATGGGCAAGAGCAAGACTTCTTC 138  
|||||  
Db 275 GGGCGTTGGCAAGTCTGCTGCTGTGCGCTTTAGCGAAGACTCATTC 323  
|||||

RESULT 13  
BG280615/c  
LOCUS 535 bp mRNA linear EST 21-FEB-2001  
DEFINITION c5blinp.fi Neurospora crassa sexual cDNA library, Uni-zap vector  
system Neurospora crassa cDNA clone c5blnp 3', mRNA sequence.  
ACCESSION BG280615  
VERSION BG280615.1 GI:13079197  
KEYWORDS EST.  
SOURCE Neurospora crassa.  
ORGANISM Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariales; Sordariaceae; Neurospora.

REFERENCE 1 (bases 1 to 535)  
AUTHORS Kupfer,D., Lai,H., Nelson,M. and Roe,B.  
TITLE ESTs from a Neurospora crassa Sexual cDNA Library  
JOURNAL Unpublished (2001)  
COMMENT Other ESTs: c5blinp.r1  
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu  
Department of Chemistry and Biochemistry  
Advanced Center for Genome Technology, University of Oklahoma  
620 Parrington Oval, Norman, OK 73019, USA  
Tel: 405 325 4912  
Fax: 405 325 7762  
Email: broe@ou.edu

FEATURES  
source  
1..535  
/organism="Neurospora crassa"  
/strain="wild type"  
/db\_xref="taxon:5141"  
/clone="c5blinp"  
/clone lib="Neurospora crassa sexual cDNA library, Uni-zap  
vector system"  
/tissue\_type="perithecia (fruiting bodies)"  
/dev\_stage="sexual"  
/lab\_host="E. coli strain SOLR"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript  
; 3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 118 a 130 c 143 g 144 t  
ORIGIN

Query Match 10.7%; Score 37; DB 12; Length 535;  
Best Local Similarity 58.7%; Pred. No. 4.6;

Matches 64; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 30 CCTCAAGAGCTCAACCGTATTCGACGGGCTGCTGAGGTTTGCTTCTCCAGGACTC 89  
 Db 388 CATCATGTCGAGTAAACAGGAATTACGATTTCTTGATCAAGCTGCTCTTGATCGGAGACTC 329

QY 90 CGGTCAACGGCAAGTCCAGCTCCAAAATGGGCAAGCGAAGACTTCTTTC 138  
 Db 328 GGGCGTTGGCAAGTCTCTGCTCTGTTGGGCTTTAGCGAAGACTCATTC 280

RESULT 14  
 BJJ268837

LOCUS BJJ268837 565 bp mRNA linear EST 09-APR-2002

DEFINITION aestivum cDNA clone whoh18024 5', mRNA sequence.

ACCESSION BJJ268837

VERSION BJJ268837.1 GI:20088745

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

AUTHORS Ogihara, Y. and Murai, K.

TITLE Expressed genes in triticum aestivum

JOURNAL Unpublished (2002)

COMMENT Contact: Tadasu Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshini@genes.nig.ac.jp.

## FEATURES

source

1. 565  
 /organism="Triticum aestivum"  
 /cultivar="Chinese Spring"  
 /db\_xref="taxon:4565"  
 /clone\_lib="whoh18024"  
 /tissue\_type="pistil at heading date"  
 /dev\_stage="Peekes' scale 10.5"  
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;  
 Site 1: EcoRI; Site 2: XhoI; Plants were grown under  
 hydroponic conditions at UC Davis, salt stressed for 12  
 hours, and for 7 days, then dissected and frozen (Akhunov  
 in J. Dvorak Lab). Total RNA was prepared from sheath  
 tissue, equal quantities of RNA were pooled from the two  
 samples, polyA was purified from the pooled RNA, a cDNA  
 library was made, and the cDNA clones were in vivo  
 excised to give phagescript phagemids in the TJ Close lab  
 at the University of California, Riverside (Akhunov, Chin  
 , Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).  
 Plasmid DNA preparations and DNA sequencing were  
 performed in the OD Anderson lab (all other authors)."

BASE COUNT 127 a 208 c 146 g 81 t 3 others

ORIGIN

Query Match 10.7%; Score 36.8; DB 13; Length 565;  
 Best Local Similarity 46.7%; Pred. No. 5.4; Indels 0; Gaps 0;  
 Matches 113; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 23 CTTCCCACTCAAGAGTCAACCGTATTCGACGGGCTGCTGAGGTTTGCTTCTCCC 82  
 Db 104 CGCGCGAGTCCACCTNGGACCAAGAACTGCGACTTCCAGATGAGCGCTACGCTACA 163

QY 83 AGGACTCCGTCACGGCAAGTCCAGCTCCAAAATGGGCAAGCGAAGACTTCTTCTACA 142  
 Db 164 AGCGCGGCTCCGACGGCATCTTACATCATCATCTGGGCAAGACATGGGGAAGCTCCAGC 223

QY 143 TCAAGGTCAAGCAGAAAGCCCGGGGCTCCCATCACCAGATTACACGGCACTGTCCAGCG 202

Db 224 TCGCGCGAGGGTTATCGTCCCATCGAAGACCCCGAGGACATCATCTGTCAGTCCGCC 283

QY 203 GGGATTTAGGGGGCCAGGCTGGGGTGGACTTTGACCGTCCACACCCCGCCCAAGCTCATGC 262

Db 284 GCCCTACGGCCAAAGCGCGCTCTCAAGTTTCGGCGAGCACACAGGGGCCAACGCCATCG 343

QY 263 CC 264

Db 344 CC 345

## RESULT 15

LOCUS AL594733

DEFINITION AL594733 XGC-gastrula Silurana tropicalis cDNA clone Tgas001a14 5', mRNA sequence.

ACCESSION AL594733

VERSION AL594733.1 GI:15006794

KEYWORDS EST.

SOURCE western clawed frog.

ORGANISM Silurana tropicalis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Silurana.

AUTHORS 1 (bases 1 to 601)  
 Huckle, E., Taylor, R., McMurray, A., Ashurst, J.L., Zorn, A.M. and Rogers, J.

TITLE Sanger Xenopus tropicalis EST project 2001

JOURNAL Unpublished (2001)

COMMENT Contact: Huckle E  
 Sanger Centre  
 Hinxton, Cambridgeshire, CB10 1SA, UK  
 Email: trop@sanger.ac.uk  
 Sanger Xenopus tropicalis EST project 2001  
 TROPICALIS SEQUENCE ID: Tgas001a14.sp6  
 Sequencing primer: SP6  
 This sequence is from a Xenopus Gene Collection (XGC) library  
 constructed by Aaron M. Zorn.

## FEATURES

source

1. 601  
 /organism="Silurana tropicalis"  
 /db\_xref="taxon:8364"  
 /clone\_lib="Tgas001a14"  
 /dev\_stage="gastrula (stages 10.5-13 mixed)"  
 /lab\_host="Escherichia coli XL1-blue"  
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
 was oligo dt primed from sug of poly A+ RNA from stages  
 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated  
 into pCS107 with EcoRI at the 5' end and NotI at the 3'  
 end."

BASE COUNT 165 a 157 c 180 g 98 t 1 others

ORIGIN

Query Match 10.6%; Score 36.6; DB 9; Length 601;  
 Best Local Similarity 63.5%; Pred. No. 6.3; Indels 3; Gaps 1;  
 Matches 73; Conservative 0; Mismatches 39; Indels 3; Gaps 1;

QY 46 CGGTATTCCGACGGGCTGCTGAGGTTTGCGTCTCCAGGACTCCGGTCAAGGCAAGTCC 105

Db 236 CTGTGGTACTTCGAGCTGCTGAAGTTCCACCATCAGCAGGAGCGCAG---CGGCAACGCC 292

QY 106 AGCTCCAAAATGGGCAAGAGCGAAGACTTCTTCTACATCAAGGTCAGCCAGAAAG 160

Db 293 ACCTCCAACCTGGGGGAGAGCGAGGACATCATCTTCCGACGTCGGGGAGGAGG 347

Search completed: June 15, 2003, 03:41:48  
 Job time : 484.339 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 2092.33 Seconds  
(without alignments)  
11183.052 Million cell updates/sec

Title: US-09-513-888C-1\_COPY\_1707\_2510

Perfect score: 804

Sequence: 1 gccatctgactctccccc.....tttggcgccagatcccccct 804

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.on.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg\_rod.\*

36: em.htg\_mam.\*

37: em.htg\_vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	804	100.0	9108	9	AF123653	AF123653 Homo sapi
2	802.4	99.8	173264	2	AC025853	AC025853 Homo sapi
3	746.2	92.8	1515	9	AF123656	AF123656 Homo sapi
4	746.2	92.8	1614	9	AF123655	AF123655 Homo sapi
5	746.2	92.8	1692	9	AF123657	AF123657 Homo sapi
6	746.2	92.8	1722	9	AF123658	AF123658 Homo sapi
7	746.2	92.8	5492	9	AF123659	AF123659 Homo sapi
8	542.2	67.4	227884	2	AC114995	AC114995 Mus muscu
9	542.2	67.4	263546	2	AC099416	AC099416 Mus muscu
10	233.2	29.0	191210	2	AC108987	AC108987 Rattus no
11	157.6	19.6	633	9	AF123654	AF123654 Homo sapi
12	152	18.9	125856	9	HS1187M17	AL121891 Human DNA
13	151	18.8	5257	9	AB011124	AB011124 Homo sapi
14	148.2	18.4	205581	10	AL731707	AL731707 Mouse DNA
15	148.2	18.4	220355	2	AL772162	AL772162 Mus muscu
16	136.6	17.0	2416	10	RNO278801	AC129762 Rattus no
17	130	16.2	159724	2	AC129762	AY029201 Homo sapi
18	125.4	15.6	1935	9	AY029201	AB046013 Macaca fa
19	125.4	15.6	2099	9	AB046013	AL834338 Homo sapi
20	125.4	15.6	2275	9	HS0805394	BC006212 Homo sapi
21	125.4	15.6	2766	9	BC006212	AB058716 Homo sapi
22	125.4	15.6	5733	9	AB058716	BC014695 Mus muscu
23	113.8	14.2	2767	10	BC014695	AX400518 Sequence
24	112.4	14.0	561	6	AX400518	AL133215 Human DNA
25	94	11.7	181086	9	AL133215	AK097997 Homo sapi
26	92.4	11.5	2308	9	AK097997	BC005855 Homo sapi
27	87.8	10.9	1612	9	BC005855	AC106674 Rattus no
28	86.4	10.7	180418	2	AC106674	AL833803 Mus muscu
29	85.4	10.6	225045	2	AL833803	AC099443 Rattus no
30	84.8	10.5	177035	2	AC099443	AC122329 Mus muscu
31	84.6	10.5	167473	2	AC122329	AC116700 Mus muscu
32	84.6	10.5	168210	2	AC116700	AC119801 Mus muscu
33	84.6	10.5	251958	2	AC119801	AC121787 Mus muscu
34	84.6	10.5	325808	2	AC121787	AC130883 Rattus no
35	84.4	10.5	157147	2	AC130883	Continuation (4 of
36	83.8	10.4	103729	2	AC096620	AC105485 Rattus no
37	83.8	10.4	270700	2	AC105485	AL606742 Mouse DNA
38	83.2	10.3	186559	10	AL606742	AC090533 Mus muscu
39	83	10.3	272545	2	AC090533	U02079 Epstein-Bar
40	82.8	10.3	1150	14	HS4ULIR3	AX107940 Sequence
41	82.8	10.3	1926	6	AX107940	AR108994 Sequence
42	82.8	10.3	2580	6	AR108994	AR083151 Sequence
43	82.8	10.3	5452	6	AR083151	U02454 Cloning vec
44	82.8	10.3	5452	12	U02454	A92665 Sequence 1
45	82.8	10.3	9600	6	A92665	

#### ALIGNMENTS

RESULT 1  
AF123653  
LOCUS  
DEFINITION Homo sapiens FEZ1 (FEZ1) gene, complete cds.  
ACCESSION AF123653  
VERSION AF123653.1 GI:4572463  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 9108)  
REFERENCE  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE The FEZ1 gene at chromosome 9p22 encodes a leucine-zipper protein,





Institute; 233S 10th street, Philadelphia, PA 19107, USA									
FEATURES	Location/Qualifiers								
source	1. .1515								
	/organism="Homo sapiens"								
	/db xref="taxon:9606"								
	/chromosome="8"								
	/map="8p22"								
	/clone="D14"								
gene	1. .1515								
	/gene="FEZ1"								
CDS	1. .1515								
	/gene="FEZ1"								
	/note="alternatively spliced"								
	/codon_start=1								
	/product="FEZ1"								
	/protein_id="A4D23837.1"								
	/db xref="GI:4572470"								
	/translation="MGSVSLISGHSFHKPCASOYKLRKSHLKLNRYSDGLLRP								
	FRSDDGHKSSKMGKSDRDFYIKVQKARQSHHEDYTLSSGDLGGAGQVDPDRST								
	PPKLMFSPNQLMGSEBKGAVRPAFKVILPRSGAILHSSPESASHQHPAPDPKPEQ								
	ELKPGICSGALSDGRNSWSLSLTHSTSSYQLDPLVTVPVPTSRFGSAHNITQIV								
	LDSNNMISLKALISFDGGKSLGHSNKADKSPSCVRSPISTDECSIQLSQKLLERGA								
	LOKLORFEEKELASLAEYERPRRCRDELEGPPEKGNKIKOASOSKROAOOVLHQ								
	VLOLOEQKRLQLESLAKMKDOLLTKLRSVEREKTSPGALFEETQWEVCOKSGRIS								
	LLKQKJESQTEVNAKASILGLKQIKETRGLKLEGLRLTQDLEGALRTKGLLEBVC								
	ENELQSYVAMYQRNRLKALQQLARGDSAGSPLFVLDLEGADIPYEDIAETI"								
BASE COUNT	336 a	483 c	482 g	214 t					
ORIGIN									
Query Match	92.8%;	Score	746.2;	DB	9;	Length	1515;		
Best Local Similarity	99.6%;	Pred. No.	1e-111;	3;	Indels	0;	Gaps	0;	
Matches	748;	Conservative	0;	Mismatches	3;				
QY	1	GCCATCTGCATCTCTCCCGGAGAGTGC	CAGCAC	CACGACTGC	ACCCGCGCCCTCCAGAC	60			
Db	406	GCCATCTGCATCTCTCCCGGAGAGTGC	CAGCAC	CACGACTGC	ACCCGCGCCCTCCAGAC	465			
QY	61	AAGCCCAAGAGCAGAGCTGAAGCTTGG	CGCTGTGCTCTGGGGCGCTGT	CAGACTCGGC	120				
Db	466	AAGCCCAAGAGCAGAGCTGAAGCTTGG	CGCTGTGCTCTGGGGCGCTGT	CAGACTCGGC	525				
QY	121	CGGAATCTCATGTCCAGCCTGCCACAC	CAGACAC	CAGCAGCAGCTAC	CAGCTGGACCCG	180			
Db	526	CGGAATCTCATGTCCAGCCTGCCACAC	CAGACAC	CAGCAGCAGCTAC	CAGCTGGACCCG	585			
QY	181	CTGGTCACACCCCGTGGGACCCACAAG	CCGTTTTTTGGGGGCTCCG	CCCCACAACATCACCCAG	240				
Db	586	CTGGTCACACCCCGTGGGACCCACAAG	CCGTTTTTTGGGGGCTCCG	CCCCACAACATCACCCAG	645				
QY	241	GGCATCGTCTCCAGGACAGCAACATGAT	GAGCCTGAAGGCTCTGTCTTCTCG	ACCGGA	300				
Db	646	GGCATCGTCTCCAGGACAGCAACATGAT	GAGCCTGAAGGCTCTGTCTTCTCG	ACCGGA	705				
QY	301	GGTAGCAAGCTGGGCACACTCGAACAG	GCGACAGCAAGGGCCCTCTGTGTG	CCGCTCCCCC	360				
Db	706	GGTAGCAAGCTGGGCACACTCGAACAG	GCGACAGCAAGGGCCCTCTGTGTG	CCGCTCCCCC	765				
QY	361	ATCTCCACGGACGAGTGCAGCATCCAG	GAGCTGGAA	CAGAAGCTGTTTGGAGAGGGAGGC	420				
Db	766	ATCTCCACGGACGAGTGCAGCATCCAG	GAGCTGGAG	CAGAAGCTGTTTGGAGAGGGAGGC	825				
QY	421	GCCCTCAGAAGCTGCAGCGCAGCTTTG	AGGAGAGAGGAGCTTGCCTCCAG	CGCTTGGCCCTAC	480				
Db	826	GCCCTCAGAAGCTGCAGCGCAGCTTTG	AGGAGAGAGGAGCTTGCCTCCAG	CGCTTGGCCCTAC	885				
QY	481	GAGGAGCGCGCGCGCTCGAGGACGAG	CTGGAGGGCCCGGAGCCCAAGCGGCAAC	540					
Db	886	GAGGAGCGCGCGCGCTCGAGGACGAG	CTGGAGGGCCCGGAGCCCAAGCGGCAAC	945					
QY	541	AAGCTCAAGCAGGCCTCGCAGAGAGCC	AGCGCGCAGCAGGCTCTGCACCTG	CAGGTA	600				
Db	946	AAGCTCAAGCAGGCCTCGCAGAGAGCC	AGCGCGCAGCAGGCTCTGCACCTG	CAGGTA	1005				







and its expression is altered in multiple human tumors  
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
99199287  
10097140  
2 (bases 1 to 5492)  
Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,  
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
Direct Submission  
Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
Institute, 2335 10th street, Philadelphia, PA 19107, USA  
Location/Qualifiers  
1. 5492  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
1. 5492  
/gene="FEZ1"  
112..1902  
/gene="FEZ1"  
/codon\_start=1  
/product="FEZ1"  
/protein\_id="AAD23840.1"  
/db\_xref="GI:4572476"  
/translation="MGSVSSLSIGHSFHSHKCRASQYKLRKSHLKLNRVSDGLLRP  
GFSQDSHGKSSKMGKSEDFYIKVQKARGSHHPDYIALSSDGLGGAGVDPDPST  
PKLPMTSNQIEMSEKGAFTKPVLPKRGAILHSSPESASHQHPADPKPEQ  
ELKPLCSGALSIDGRMSLPHSTSSSYQDLPVTPVGTFRFGSAHNTQGI  
LQDNMSLKALSDGSKLGHKADKPCSVRSPISTDECISIOLEOKLAREGA  
LQKQRFEEELASSLAYERPRCRDELEGPEPKGNKLKQASQSQRAQVHLQ  
VLQQKQKQRQNELSLMKQDLETKLRSYREKTSFGLPABETQWECQKSGEIS  
LLKQOLKESEVNAKASEILGLKAQKDTGRGLELELRTQDLEGALRTKGLELVC  
ENLQKKEAELEKRNLEQLQELRAQALARDMGPTTFDPFALQRELERLR  
AELREERQGHDMSSGFOHERLVNKEKEKVIQKQLOQSYVAMYQNRQLEKALQ  
LARDGSAPELVELEGADIPYDIATFI"  
BASE COUNT 1137 a 1704 c 1565 g 1086 t

Query Match 92.8%; Score 746.2; DB 9; Length 5492;  
Best Local Similarity 99.6%; Pred. No. 8.2e-112;  
Matches 748; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTGCCAGCACAGCTGCACCCCGCCCTCCAGAC 60  
Db 517 GCCATCTGCACTCTCCCGGAGAGTGCCAGCACAGCTGCACCCCGCCCTCCAGAC 576  
Qy 61 AAGCCCAAGGAGGAGGAGCTGAAGCTGGCTGTCTCTGGGGGCTGTCTGAGCTCCGC 120  
Db 577 AAGCCCAAGGAGGAGGAGCTGAAGCTGGCTGTCTCTGGGGGCTGTCTGAGCTCCGC 636  
Qy 121 CGGAATCTCATGTCTCAGCTGCCACACACAGCACAGCAGCTACCACTGAGCCCG 180  
Db 637 CGGAATCTCATGTCTCAGCTGCCACACACAGCACAGCAGCTACCACTGAGCCCG 696  
Qy 181 CTGCTACACCCGTGGGAGCCACAGCGTTTGGGGGCTCCGCCCAACATCAACCCAG 240  
Db 697 CTGCTACACCCGTGGGAGCCACAGCGTTTGGGGGCTCCGCCCAACATCAACCCAG 756  
Qy 241 GGCATCTGCTCCAGGACGACATGATGAGCTGAGGCTGTCTCTTCTCCGACGGA 300  
Db 757 GGCATCTGCTCCAGGACGACATGATGAGCTGAGGCTGTCTCTTCTCCGACGGA 816  
Qy 301 GGTAGCAGCTGGGCACTCGAAAGGACGACAGGCGCCCTCGTGTGCTCCGCTCCCCC 360  
Db 817 GGTAGCAGCTGGGCACTCGAAAGGACGACAGGCGCCCTCGTGTGCTCCGCTCCCCC 876  
Qy 361 ATCTCCAGGACGAGTGAGCATCCAGAGCTGGAAACAGAGCTTTGGAGGAGGAGGC 420  
Db 877 ATCTCCAGGACGAGTGAGCATCCAGAGCTGGAAACAGAGCTTTGGAGGAGGAGGC 936  
Qy 421 GCGCTCCAGAGCTGCAGCGAGCTTTGAGGAGGAGGCTTGCCTCCAGGCTGCCTAC 480  
Db 937 GCGCTCCAGAGCTGCAGCGAGCTTTGAGGAGGAGGAGCTTGCCTCCAGGCTGCCTAC 996

Qy 481 GAGGAGCGCGCGCGCTGCAGGAGAGCGAGCTGAGGCGCGCGAGCCCAAGGGCGCAAC 540  
Db 997 GAGGAGCGCGCGCGCTGCAGGAGAGCGAGCTGAGGCGCGCGAGCCCAAGGGCGCAAC 1056  
Qy 541 AAGCTCAAGCAGGCTGCAGAGAGGAGCGCGCGAGCGCTCTGCACCTGCAGGTA 600  
Db 1057 AAGCTCAAGCAGGCTGCAGAGAGGAGCGCGCGAGCGCTCTGCACCTGCAGGTA 1116  
Qy 601 CTGAGGCTTCAGCAGGAGAGCGCGAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGGAG 660  
Db 1117 CTGAGGCTTCAGCAGGAGAGCGCGAGCTCCGCGAGGAGCTCGAGAGCTCATGAAGGAG 1176  
Qy 661 CAGGACCTGCTGGAGACCAAGCTCAGGTCTCTACAGAGGAGAGAGACAGCTTCGCGCCC 720  
Db 1177 CAGGACCTGCTGGAGACCAAGCTCAGGTCTCTACAGAGGAGAGAGACAGCTTCGCGCCC 1236  
Qy 721 GCGCTGGAGGAGACCCAGTGGAGGTGAGGC 751  
Db 1237 GCGCTGGAGGAGACCCAGTGGAGGTGAGGC 1267

RESULT 8  
AC114995 227884 bp DNA linear HTG 26-JUN-2002  
LOCUS Mus musculus clone RP23-171P5, WORKING DRAFT SEQUENCE, 5 ordered  
DEFINITION pieces.  
AC114995  
AC114995.3 GI:21592059  
VERSION HTG; HTGS\_PHASE3; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS house mouse.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 227884)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Mus musculus, clone RP23-171P5  
Unpublished  
2 (bases 1 to 227884)  
Anderson, S., Barna, N., Nuebaum, C., Lander, E., Ali, A., Allen, N.,  
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A.,  
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Farro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Jones, C.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,  
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,  
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,  
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 227884)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,  
Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collamore, A.,  
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Farro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,





## RESULT 10

AC108987

LOCUS

DEFINITION

AC108987 . 191210 bp DNA linear HTG 13-JUL-2002  
 Rattus norvegicus clone CH230-115K1, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 78 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AC108987.3 GI:21737647  
 HTG: HTGS\_PHASE1.  
 Norway rat.  
 Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

## REFERENCE

AUTHORS

1 (bases 1 to 191210)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
 Albrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,  
 Barbala, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,  
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,  
 Bunay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Gartner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,  
 Homs, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maneshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenko, S., Ogih, M., Okwuonu, G.,  
 Oreguay, N., Oviedo, R., Pace, A., Payton, B., Peary, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
 Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,  
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vaequez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Direct Submission  
 Unpublished  
 2 (bases 1 to 191210)  
 Worley, K.C.

Direct Submission

Submitted (03-FEB-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 191210)

Worley, K.C.

Direct Submission

Submitted (13-JUL-2002)

Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 12, 2002 this sequence version replaced gi:18846600.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GPMW

Center clone name: CH230-115K1

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 115719 bases at least Q40

Consensus quality: 122885 bases at least Q30

Consensus quality: 129025 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 78 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1  
 1041: contig of 1041 bp in length  
 1141: gap of unknown length  
 1142: contig of 1121 bp in length  
 2262: gap of unknown length  
 2362: contig of 1176 bp in length  
 3538: gap of unknown length  
 3638: gap of unknown length  
 4775: contig of 1137 bp in length  
 4875: gap of unknown length  
 6050: contig of 1175 bp in length  
 6150: gap of unknown length  
 7660: contig of 1510 bp in length  
 7760: gap of unknown length  
 9331: contig of 1571 bp in length  
 9431: gap of unknown length  
 11027: contig of 1596 bp in length  
 11127: gap of unknown length  
 12869: contig of 1642 bp in length  
 12869: gap of unknown length  
 14092: contig of 1223 bp in length  
 14192: gap of unknown length  
 15228: contig of 1036 bp in length  
 15328: gap of unknown length  
 16709: contig of 1381 bp in length  
 16809: gap of unknown length  
 18195: contig of 1386 bp in length  
 18295: gap of unknown length  
 19412: contig of 1117 bp in length  
 19512: gap of unknown length  
 20959: contig of 1447 bp in length  
 21059: gap of unknown length  
 22161: contig of 1102 bp in length  
 22261: gap of unknown length  
 23587: contig of 1326 bp in length  
 23687: gap of unknown length  
 25231: contig of 1544 bp in length  
 25331: gap of unknown length  
 27135: contig of 1804 bp in length  
 27235: gap of unknown length  
 29175: contig of 1940 bp in length  
 29275: gap of unknown length  
 30846: contig of 1571 bp in length  
 30946: gap of unknown length  
 32090: contig of 1144 bp in length  
 32190: gap of unknown length  
 34039: contig of 1849 bp in length  
 34139: gap of unknown length  
 35408: contig of 1269 bp in length  
 35508: gap of unknown length  
 37018: contig of 1510 bp in length

* 37019	37118: gap of unknown length	86417	AAGCTACAGCGCAGTTTCGATGAGAAGTAGTTTACCTCTGTCCAGACTTTCGAGGAGCGA	86476	
* 37119	38260: contig of 1142 bp in length	Qy	490	CCGCGCGCTGCAGGAGCAGAGCTGAGGGCCGCGAGCCCAAGGCGGCAACAAGCTCAAG	549
* 38261	38360: gap of unknown length	Db	86477	CCTCGCTGGACTAGACAGCACTGGAGTGCCTGCAATAA-----GAGCAAGCTGAAG	86530
* 40084	40083: contig of 1723 bp in length	Qy	550	CAGCCTCGCAGAGAGCCAGCGCGCAGAGCTCTGACCTTCCAGGTACTCCAGCTT	609
* 40184	41200: contig of 1017 bp in length	Db	86531	CCATCATCTCAGAAAGATCCAGCGCACCCAGCAGTACTCCAGCTGCAGGTGTTCCAGTTG	86590
* 41201	41300: gap of unknown length	Qy	610	CAGCAGGAGAGCGGCGAGCTCCGCGAGAGCTCCAGAGCTCATGAGGACGAGACCTG	669
* 43365	43365: contig of 2065 bp in length	Db	86591	CAGCAGGAGAGTCCGCGAGCTGAGCGAGGCTAGAGAGCTTAAAGAGGACGAGACCTG	86650
* 43366	43465: gap of unknown length	Qy	670	CTGAGAGACCAAGCTCAGGTCTACGAGAGGAGAGACACAGCTTCGCGCCCGCTGGAG	729
* 43442	45441: gap of unknown length	Db	86651	CTGGAGACCAAGCTCAGGTCTGAGCGGAGGAGATAGCCATCTTCGCCCATGCACTGGAG	86710
* 43443	45442: contig of 1901 bp in length	Qy	730	GAGACCCAGTGGGAGGTGAGGCCACACAGGGGC	761
* 43444	45443: gap of unknown length	Db	86711	GAGACCCAAATGGGAGGTAAAGACCCCACTAGTC	86742
* 45444	45987: contig of 1544 bp in length	RESULT 11			
* 46988	47087: gap of unknown length	AF123654			
* 47088	49404: contig of 2317 bp in length	LOCUS			
* 49045	49504: gap of unknown length	DEFINITION	Homo sapiens clone E264162 FEZ1 (FEZ1) mRNA, linear	PRI 07-APR-1999	
* 50824	50923: gap of unknown length	complete cds.			
* 50924	52451: contig of 1528 bp in length	ACCESSION	AF123654		
* 52452	52551: gap of unknown length	VERSION	AF123654.1	GI:4572465	
* 52552	54441: contig of 1890 bp in length	KEYWORDS			
* 54442	54441: gap of unknown length	SOURCE			
* 54442	56442: contig of 1901 bp in length	ORGANISM	Homo sapiens.		
* 56443	56542: gap of unknown length	REFERENCE			
* 56543	57449: contig of 1207 bp in length	AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,		
* 57449	57849: gap of unknown length	TITLE	The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,		
* 57850	60138: contig of 2289 bp in length	JOURNAL	and its expression is altered in multiple human tumors		
* 60139	60238: gap of unknown length	MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)		
* 60239	62467: contig of 2229 bp in length	PUBMED	99199287		
* 62468	62567: gap of unknown length	REFERENCE	2 (bases 1 to 633)		
* 62568	64364: contig of 1797 bp in length	AUTHORS	Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,		
* 64365	64464: gap of unknown length	TITLE	Direct Submission		
* 64465	66889: contig of 2425 bp in length	JOURNAL	Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer		
* 66890	66989: gap of unknown length	MEDLINE	Institute, 233S 10th street, Philadelphia, PA 19107, USA		
* 66990	68330: contig of 1841 bp in length	PUBMED	10097140		
* 68331	68930: gap of unknown length	FEATURES			
* 68931	71226: contig of 2296 bp in length	source			
* 71227	71326: gap of unknown length	/organism="Homo sapiens"			
* 71327	73112: contig of 2386 bp in length	/db_xref="taxon:9606"			
* 73113	73812: gap of unknown length	/chromosome="8"			
* 73813	76871: contig of 3059 bp in length	/map="8p22"			
* 76872	76871: gap of unknown length	/clones="E264162"			
* 76972	79844: contig of 2873 bp in length	1..633			
* 79845	81259: contig of 1315 bp in length	/gene="FEZ1"			
* 81260	81359: gap of unknown length	1..633			
* 81360	83690: contig of 2331 bp in length	/gene="FEZ1"			
* 83691	83790: gap of unknown length	/note="alternatively spliced"			
* 83791	86195: contig of 2405 bp in length	/codon_start=1			
* 86196	86295: gap of unknown length	/product="FEZ1"			
* 86296	88065: contig of 1770 bp in length	/protein_id="A4572466"			
* 88066	88165: gap of unknown length	/db_xref="GI:4572466"			
* 88166	90384: contig of 2219 bp in length	1..633			
* 90385	90484: gap of unknown length	/gene="FEZ1"			
* 90485	91999: contig of 1515 bp in length	1..633			
* 92000	92099: gap of unknown length	/note="alternatively spliced"			
* 92100	94583: contig of 2484 bp in length	/codon_start=1			
* 94584	94683: gap of unknown length	/product="FEZ1"			
* 94684	98297: contig of 3614 bp in length	/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"			
		/codon_start=1			
		/product="FEZ1"			
		/protein_id="A4572466"			
		/db_xref="GI:4572466"			
		1..633			
		/gene="FEZ1"			
		1..633			
		/note="alternatively spliced"	</		









Db 183362 AGCAGGAGGTGGCAGCTCTGGCGGCCAGCCTTGAGCAGAGCAGCAGCCTGGCCCCAGG 183303  
Qy 476 CCTACGAGGAGCGCGCGCGCTGCAGGAGAGCTGGAGGGCCGCGAGCCCAAGGCG 535  
Db 183302 TGTGGAGAGCGGAGAGGATGGGAGCGGAGCTAGCCAGTCTCGGAGGGCTGCA 183243  
Qy 536 GCAACAAGCTCAAGCAGCCTGCGAGAGAGCCAGCGCGCAGCAGCTCTGCACTTCG 595  
Db 183242 GTGGGAAGCTGCAGCAGGTGGCGCCCGCTGCCAGCGTCCAGCAGGCGCTACAGTGC 183183  
Qy 596 AGGTACTGCAGTTCAGCAGGAGAGCGCGAGCTCGGAGAGAGCTCGAGGCCTCATGA 655  
Db 183182 AGGTCTCGGTTACAGCAGGACAAAGACTGCAGGAGAGCGCGCCAGCTGTATA 183123  
Qy 656 AGGACGAGCCTCTGAGAGCAAGCTCAGGTCTCTAGAGAGGAGAGACCAAGCTTCG 715  
Db 183122 GCACAGGAGAGCTGGAGGACAGGTGGCGCTGTCTAGAGAGGAGAGCGGCACCTTC 183063  
Qy 716 GCCCGCGCTGAGAGAGCCAGTGGGAGGTGAGGCC 752  
Db 183062 TGCCCCGATGAGGAAACTAAGTGGAGGTGCGGCG 183026

## RESULT 15

AL772162/c  
LOCUS AL772162 220355 bp DNA linear HTG 17-AUG-2002  
DEFINITION Mus musculus chromosome 2 clone RP23-175J8, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 4 unordered pieces.

ACCESSION AL772162

VERSION AL772162.3 GI:22415967

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.

SOURCE house mouse.

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

A Almeida, J.

Direct Submission

Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Aug 21, 2002 this sequence version replaced gi:22002755.

## COMMENT

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: bm175J8

----- Summary Statistics

Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 91% of reads  
Chemistry: Dye-primer Big Dye; 0% of reads  
Consensus quality: 220027 bases at least Q40  
Consensus quality: 220037 bases at least Q30  
Consensus quality: 220045 bases at least Q20  
Insert size: 220055; sum-of-contigs  
Insert size: 215691; 7.3% error; agarose-fp  
Quality coverage: 8.70x in Q20 bases; sum-of-contigs Quality  
coverage: 8.99x in Q20 bases; agarose-fp

----- NOTE: This is a 'working draft' sequence. It currently

\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 152569: contig of 152569 bp in length

\* 152570 152669: gap of 100 bp

\* 152670 197614: contig of 44945 bp in length

\* 197615 197714: gap of 100 bp  
\* 197715 216175: contig of 18461 bp in length  
\* 216176 216275: gap of 100 bp  
\* 216276 220355: contig of 4080 bp in length.  
FEATURES  
source  
1..220355  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="2"  
/clone="RP23-175J8"  
/clone\_lib="RPC1-23"  
1..152569  
/note="assembly fragment:02397  
fragment\_chain:1  
clone\_end:T7  
vector\_side:right"  
152670..197614  
/note="assembly fragment:03756  
fragment\_chain:1"  
197715..216175  
/note="assembly fragment:02402  
fragment\_chain:1"  
216276..220355  
/note="assembly fragment:03760  
fragment\_chain:1"  
BASE COUNT 58827 a 52680 c 52199 g 56349 t 300 others  
ORIGIN  
Query\_Match 18.4%; Score 148.2; DB 2; Length 220355;  
Best Local Similarity 53.7%; Pred. No. 4.4e-15;  
Matches 374; Conservative 0; Mismatches 293; Indels 30; Gaps 2;  
Qy 86 CTGGCGCTGTCTCTGGGGCGCTGTCCAGACTCCGCGCGGAATCTCATGTCACAGCTTGCCCA 145  
Db 100302 CTGTGGAGTGGGGGGCGCTCTCAGACTCAGCGCGGAATCTACTCACAAGCTTGCCCA 100243  
Qy 146 CACACAGCACCAGCAGCAGCTACCACTGGACCCGCTGGTGCACACCCGTTGGAGCCCA 205  
Db 100242 CCTATAGCTCCAGCTATATAGCCAGCATCTGGCACCCCTCAGTGTCTCCACAGCATTATCA 100183  
Qy 206 GCCGTTTGGGGCTCGGCCCAACAT-----CACCAGG 241  
Db 100182 ACCGTATTTGGCACCGCTGGCTATAGTGTGAGCAGCGGTGGGGGTTCAGGTACACAGG 100123  
Qy 242 GCATCGTCTCCAGCAGCAACATGATGACCTTGAAGGCTCTGTCTCTCCGACGGAG 301  
Db 100122 ATCTGGGACCTCTGACAGTGGCGGGCTTCCAGTAAAGTGGTGTCTATCATTCATCGG 100063  
Qy 302 G-----TAGCAAGCTGGGCCACTCGAAACAGGCGAGCAAGGGCCCTCGTGTGTCCGCT 355  
Db 100062 GGCGGTCAAGGCCACTCGGATCCGGGAGGGCGGAAATGGGGCTTGCCATTTGCAGCCT 100003  
Qy 356 CCCCCTCTCCAGCAGAGTGCAGATCCAGGAGCTGGAACAGAGCTTGTGAGAGGG 415  
Db 100002 GTTCACCACTCTCCGCCCGCTGATCTCCAGGAGCTGGAGAGCGGTGTGGGAGAGG 99943  
Qy 416 AGGGCGCTCCCAAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGG 475  
Db 99942 AGCAGGAGTGGAGCTCTGCGGGCGCAGCTTGGAGAGCAGAGCAGCGCTGGGCCAGG 99883  
Qy 476 CCTACGAGGAGCGCGCGCTGTCAGGGACGAGCTGGAGGGCCCGGAGGCCCAAGCGG 535  
Db 99882 TGTGGAGGAGCGCGCAGAGGATGGGAGCGGAGCTAGCCAGCTCCGCGAGGGCTGCA 99823  
Qy 536 GCAACAGCTCAACAGGCTCGCAGAGAGAGCCAGCGCGCAGAGGTCTCTGACCTGC 595  
Db 99822 GTGGGAAGCTGCAGCAGGTGGCGCGCTGTGCCAGCTGCCAGCAGCGCTTACAGCTGC 99763  
Qy 596 AGGTACTGCAGCTTTCAGCAGGAGAGCGGCGAGCTCCCGCAGAGCTCCAGAGCCTCATGA 655  
Db 99762 AGGTGCTGCGGTTATACAGCAGGACAAAGACTGTCAGGAGGAGCGCGCCCGCTGATAA 99703  
Qy 656 AGGAGCAGGACCTCTGTGAGACCAAGCTCAGGTCTCTACGAGAGGAGAGACCAAGCTTCG 715

Db 99702 GGCAACGGGAAGAGCTGGAGGACAAAGGTGGCCGTCTGTCAAGAGGAGCAGGCCGACTTCC 99643  
Qy 716 GCCCCCGCGCTGGAGGAGAGCCAGTGGGAGGTGAGGCC 752  
Db 99642 TGCCCCGGATGGAGGAAACTAAGTGGGAGGTCCGGGC 99606

Search completed: June 15, 2003, 02:18:36  
Job time : 2097.33 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:38:41 ; Search time 185.106 Seconds  
(without alignments)  
9781.434 Million cell updates/sec

Title: US-09-513-888c-1\_COPY\_1707\_2510

Perfect score: 804

Sequence: 1 gccatcctgactctcccc.....tttggcgccagatccccct 804

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq 101002.\*

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	804	100.0	9048	21	AAA64507
2	802.4	99.8	2411	22	AAA30637
3	802.4	99.8	2411	22	AAA28699
4	746.2	92.8	1512	21	AAA64513
5	746.2	92.8	1614	21	AAA64512
6	746.2	92.8	1692	21	AAA64514
7	746.2	92.8	1722	21	AAA64515
8	746.2	92.8	1791	21	AAA64509
9	746.2	92.8	5492	21	AAA64508

10	463.4	57.6	560	22	AAA30595	DNA encoding novel
11	463.4	57.6	560	22	AAA28127	Novel cDNA encoding
12	157.6	19.6	633	21	AAA64511	Nucleotide sequence
13	112.4	14.0	561	24	ABK62287	Rat sequence diffe
14	82.8	10.3	795	19	AAV55830	FLGA insert stabl
15	82.8	10.3	799	19	AAV55831	Nucleotide sequenc
16	82.8	10.3	1926	21	AAA50254	Epstein Barr virus
17	82.8	10.3	1926	22	AAF82902	EBV tethering prot
18	82.8	10.3	2580	21	AAA75454	Nucleotide sequenc
19	82.8	10.3	2580	24	AAI64275	Epstein-Barr virus
20	82.8	10.3	5452	20	AAV90923	Anti-sense strand
21	82.8	10.3	8705	20	AAZ23778	Vector pshuttle DN
22	82.8	10.3	9600	19	AAV21683	Vector plasmid pCM
23	82.8	10.3	10380	20	AAZ22248	Nucleotide sequenc
24	82.8	10.3	10596	14	AAQ51731	Plasmid pCisEBON f
25	82.8	10.3	10596	17	AAT40348	Plasmid pCisEBON f
26	82.8	10.3	10596	20	AAI15650	Nucleotide sequenc
27	82.8	10.3	16080	21	AAA59553	DNA clone pCEK Cl.
28	79.6	9.9	1623	24	ABQ54361	Human ovarian anti
29	77.2	9.6	1925	20	AAV90924	Epstein Barr Virus
30	77.2	9.6	5120	22	AAC84677	DNA sequence of hu
31	68.8	8.6	16442	18	AAV83006	Partial mouse WRN
32	67	8.3	1617	21	AAV15625	Human prostate can
33	67	8.3	2108	24	ABL67774	Oesophagus cancer
34	67	8.3	3489	21	AAA30290	Kaposi's sarcoma-a
35	67	8.3	3489	22	AAF82901	Nucleotide sequenc
36	67	8.3	3489	24	ABA93487	Kaposi's sarcoma-a
37	67	8.3	32207	20	AAV73805	KSHV LUR DNA (nucl
38	67	8.3	137507	19	AAV19941	KSHV long unique c
39	64.6	8.0	9551	20	AAZ22301	CDNA encoding a hu
40	62.2	7.7	1159	21	AAA59240	An EcoRI fragment
41	61.8	7.7	477	21	AAZ44307	Human SCA7 genomic
42	61.6	7.7	1037	21	AAA59242	Exons E, C and A o
43	61.6	7.7	1472	21	AAA59241	Exons D, C, B and
44	60	7.5	1416	23	AAV71051	DNA encoding novel
45	60	7.5	4720	24	AAV94782	Human DNA sequence

ALIGNMENTS

RESULT 1  
AAA64507  
ID AAA64507 standard; DNA; 9048 BP.  
AC AAA64507;  
XX  
XX  
DT 02-JAN-2001 (first entry)  
XX  
XX  
DE Nucleotide sequence comprising the human FEZ1 gene.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
OS Homo sapiens.  
XX  
XX WO2000050565-A2.  
XX  
XX 31-AUG-2000.  
XX  
XX 25-FEB-2000; 2000WO-US04950.  
XX  
XX 25-FEB-1999; 99US-0121537.  
XX  
XX (UVE-) UNIV JEFFERSON THOMAS.  
XX  
XX Croce CM, Ishii H;  
XX  
XX WPI; 2000-558396/51.  
XX







Db 1787 GCCCTCCAGAAAGTTCGACGGCAGCTTTGAGGAGAAAGGAGCTTGCCTCCAGCCTGCGCTTAC 1846  
Qy 481 GAGGAGCGCGCGCGCTGCGAGGACAGCTTGAGGGCCCGGAGCCCAAGCGCGCAAC 540  
Db 1847 GAGGAGCGCGCGCGCTGCGAGGACAGCTTGAGGGCCCGGAGCCCAAGCGCGCAAC 1906  
Qy 541 AAGCTCAAGCAGGCTTCGAGAGAGCAGCGCGCGCAGCGCTTGCACCTGCAGGTA 600  
Db 1907 AAGCTCAAGCAGGCTTCGAGAGAGCAGCGCGCGCAGCGCTTGCACCTGCAGGTA 1966  
Qy 601 CTGCAGCTTCAGCAGGAGAGCGGCGAGCTTCGGCAGGAGCTCGAGAGCTCATGAGGAG 660  
Db 1967 CTGCAGCTTCAGCAGGAGAGCGGCGAGCTTCGGCAGGAGCTCGAGAGCTCATGAGGAG 2026  
Qy 661 CAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAGACACAGCTTCGGCCCC 720  
Db 2027 CAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAGACACAGCTTCGGCCCC 2086  
Qy 721 CGCTGGAGGAGACCCAGTGGAGGTGAGGCGCACACAGGGCTCATGGGTTTGGGTGGTCA 780  
Db 2087 CGCTGGAGGAGACCCAGTGGAGGTGAGGCGCACACAGGGCTCATGGGTTTGGGTGGTCA 2146  
Qy 781 GCGGTTTGGCGCCAGTACCCCCCT 804  
Db 2147 GCGGTTTGGCGCCAGTACCCCCCT 2170

## RESULT 4

AAA64513

ID AAA64513 standard; cDNA; 1512 BP.

XX AC AAA64513;

XX DT 02-JAN-2001 (first entry)

XX DE Nucleotide sequence of truncated FEZ1 transcript D14.

XX KW Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;  
XX KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
XX KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
XX KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
XX KW tumorigenesis; tumour survival; metastasis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..1512

XX FT /\*tag= a

XX FT /product= "truncated FEZ1"

XX PN WO200050565-A2.

XX PD 31-AUG-2000.

XX PF 25-FEB-2000; 2000WO-US04950.

XX PR 25-FEB-1999; 99US-0121537.

XX PA (UYJE-) UNIV JEFFERSON THOMAS.

XX PI Croce CM, Ishii H;

XX DR WPI; 2000-558396/51.

XX DR P-PSDB; AAB08720.

XX PT New polynucleotide homologous with a portion of one strand of the human  
XX PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
XX PT cancer -  
XX PS Disclosure; Fig 5F; 255pp; English.

XX CC The present sequence encodes a truncated human FEZ1 polypeptide. The

XX CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour

CC suppressor gene, located at chromosome location 9p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX SQ Sequence 1512 BP; 335 A; 483 C; 481 G; 213 T; 0 other;

Query Match 92.8%; Score 746.2; DB 21; Length 1512;

Best Local Similarity 99.6%; Pred. No. 7.1e-132;

Matches 748; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60

Db 406 GCCATCTGCACTCTCCCGGAGAGTGCCAGCCACAGCTGCACCCCGCCCTCCAGAC 465

Qy 61 AAGCCCAAGGACGAGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTACAGTCCGGC 120

Db 466 AAGCCCAAGGACGAGAGCTGAAGCTGGCTGTCTCTGGGGCGCTGTACAGTCCGGC 525

Qy 121 CGGAACCTCCATGTCCAGCTTCCACACACAGCAGCAGCAGCTACAGCTGACCCCG 180

Db 526 CGGAACCTCCATGTCCAGCTTCCACACACAGCAGCAGCAGCTACAGCTGACCCCG 585

Qy 181 CTGTGTACACCCGTGGGACCCACAAAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 240

Db 586 CTGTGTACACCCGTGGGACCCACAAAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 645

Qy 241 GGCATCTGCTCCAGGACAGCAATGATGAGCTGAAGGCTCTGTCTTCTCGACCGA 300

Db 646 GGCATCTGCTCCAGGACAGCAATGATGAGCTGAAGGCTCTGTCTTCTCGACCGA 705

Qy 301 GGTAGCAAGCTGGGCGCACTCGAACAAGGCACACAAGGCGCCCTCGTGTGTCGGCTCC 360

Db 706 GGTAGCAAGCTGGGCGCACTCGAACAAGGCACACAAGGCGCCCTCGTGTGTCGGCTCC 765

Qy 361 ATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAACAGAGAGCTGTTGGAGAGGAGGGC 420

Db 766 ATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAACAGAGAGCTGTTGGAGAGGAGGGC 825

Qy 421 GCCCTCAGAAAGCTGCGAGCGCAGCTTTGAGGAGAAAGAGCTTGCCTCCAGCCTGCGCTTAC 480

Db 826 GCCCTCAGAAAGCTGCGAGCGCAGCTTTGAGGAGAAAGAGCTTGCCTCCAGCCTGCGCTTAC 885

Qy 481 GAGGAGCGCGCGCGCTGCAGGAGCAGCTGAGGCGCCCGGAGCCCAAGGGCGGCAAC 540

Db 886 GAGGAGCGCGCGCGCTGCAGGAGCAGCTGAGGCGCCCGGAGCCCAAGGGCGGCAAC 945

Qy 541 AAGCTCAAGCAGGCTTCGAGAGAGCAGCGCGCAGCAGAGGTCCTGACCTGCAGGTA 600

Db 946 AAGCTCAAGCAGGCTTCGAGAGAGCAGCGCGCAGCAGAGGTCCTGACCTGCAGGTA 1005

Qy 601 CTGCAGCTTCAGCAGGAGAGCGGCGAGCTCCGGCAGAGAGCTCGAGAGCTCATGAGGAG 660

Db 1006 CTGCAGCTTCAGCAGGAGAGCGGCGAGCTCCGGCAGAGAGCTCGAGAGCTCATGAGGAG 1065

Qy 661 CAGGACCTGCTGGAGAGCAAGCTCAGGTCCTTACAGAGGGGAGAGACAGCTTCGGCCCC 720

Db 1066 CAGGACCTGCTGGAGAGCAAGCTCAGGTCCTTACAGAGGGGAGAGACAGCTTCGGCCCC 1125

Qy 721 GCGCTGGAGGAGACCCAGTGGGAGGTGAGGC 751







PS	Disclosure; Fig 5H; 255pp; English.	
XX	The present sequence encodes a truncated human FEZ1 polypeptide. The	
CC	encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour	
CC	suppressor gene, located at chromosome location 8p22. Decreased	
CC	or no expression of FEZ1 is detected in a variety of cancer cells.	
CC	Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1	
CC	also interacts with tubulin, with microtubules, and with protein	
CC	EF1-gamma. Post-translational phosphorylation and dephosphorylation	
CC	modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene	
CC	expression are useful for inducing cells to proliferate. Compounds	
CC	which modulate FEZ1 association with tubulin are useful for alleviating	
CC	tubulin hyper- or hypo- polymerisation disorders, such as those	
CC	associated with aberrant initiation of mitosis, modulation of the	
CC	initiation and rate of cell proliferation and cell growth, modulation of	
CC	cell shape, cell rigidity, cell motility, rate and stage of cellular	
CC	DNA replication, intracellular distribution of organelles, metastatic	
CC	potential of cell and cellular transformation from a non-cancerous to	
CC	cancerous phenotype. Compounds which modulate FEZ1 binding and	
CC	phosphorylation are also useful for alleviating a disorder, such as	
CC	tumorigenesis, tumour survival, growth and metastasis.	
XX	Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;	
SQ		
	Query Match 92.8%; Score 746.2; DB 21; Length 1722;	
	Best Local Similarity 99.6%; Pred. No. 7.1e-132; Indels 0; Gaps 0;	
	Matches 748; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	1 GCCATCTCTGCACTCTCCCGGAGAGTGCAGCCACCACTGCTGCACCCCGCCCTCCAGAC 60	
DB		
QY	406 GCCATCTCTGCACTCTCCCGGAGAGTGCAGCCACCACTGCTGCACCCCGCCCTCCAGAC 465	
DB		
QY	61 AAGCCCAAGGACGAGCTGAAGCTGCGCTGTGCTCTGGGGCGCTGTGAGACTCCGGC 120	
DB		
QY	466 AAGCCCAAGGACGAGCTGAAGCTGCGCTGTGCTCTGGGGCGCTGTGAGACTCCGGC 525	
DB		
QY	121 CGGAATCTCATGCTCCAGCTGCCACACACAGCAGCAGCTACCACTGACGACCG 180	
DB		
QY	526 CGGAATCTCATGCTCCAGCTGCCACACACAGCAGCAGCTACCACTGACGACCG 585	
DB		
QY	181 CTGGTTCACACCCGTGGGACCCCAAGCCGTTTGGGGGCTCCGCCCAACAATCACTCCAG 240	
DB		
QY	586 CTGGTTCACACCCGTGGGACCCCAAGCCGTTTGGGGGCTCCGCCCAACAATCACTCCAG 645	
DB		
QY	241 GCATGCTCTCCAGGACGAGCAATGATGAGCTGAAGGCTCTGCTCTTCCGACCGA 300	
DB		
QY	646 GCATGCTCTCCAGGACGAGCAATGATGAGCTGAAGGCTCTGCTCTTCCGACCGA 705	
DB		
QY	301 GGTAGCAAGCTGGGCACTCGAACAAGGACAGCAAGGGGCCCTCGTGTGTCGGCTCCGCC 360	
DB		
QY	706 GGTAGCAAGCTGGGCACTCGAACAAGGACAGCAAGGGGCCCTCGTGTGTCGGCTCCGCC 765	
DB		
QY	361 ATCTCCAGGACGAGTGCAGCATCCAGGAGTGGAAAGAGCTTTGGAGAGGAGGGC 420	
DB		
QY	766 ATCTCCAGGACGAGTGCAGCATCCAGGAGTGGAAAGAGCTTTGGAGAGGAGGGC 825	
DB		
QY	421 GCCCTCCAGAACTCCAGCGGAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGCGCTAC 480	
DB		
QY	826 GCCCTCCAGAACTCCAGCGGAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGCGCTAC 885	
DB		
QY	481 GAGGAGCGCGCGCGCTGCAAGGACGAGCTGGAGGGCCCGAGCCCAAGCGCGCAAC 540	
DB		
QY	886 GAGGAGCGCGCGCGCTGCAAGGACGAGCTGGAGGGCCCGAGCCCAAGCGCGCAAC 945	
DB		
QY	541 AAGCTTCAAGCAGGCTTCGAGAAAGAGCCAGCGCGCAGAGGCTCTGCACCTGCAGGTA 600	
DB		
QY	946 AAGCTTCAAGCAGGCTTCGAGAAAGAGCCAGCGCGCAGAGGCTCTGCACCTGCAGGTA 1005f	
DB		
QY	601 CTGACGCTTCAGAGGAGGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAG 660	
DB		
QY	1006 CTGACGCTTCAGAGGAGGCGGAGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAG 1065	
DB		
QY	661 CAGGACCTCTGAGACCAAGCTCAGGCTCTTACGAGAGGAGGAGACCACTTCGCGCCC 720	
DB		

Db	1066 CAGGACCTCTGAGAGCAAGCTCAGTCTTACGAGGAGGAGAGACAGCTTCGGCCCC 1125	
QY	721 GCGCTGAGGAGAGCCAGTGGAGGTGAGGC 751	
Db	1126 GCGCTGAGGAGAGCCAGTGGAGGTGAGGC 1156	
RESULT 8		
AAA64509		
ID	AAA64509 standard; cDNA; 1791 BP.	
AC	AAA64509;	
XX	02-JAN-2001 (first entry)	
DT	cDNA sequence encoding a human FEZ1 polypeptide.	
DE	Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;	
KW	tumour proliferation; tubulin; microtubule; protein EF1-gamma;	
KW	tubulin polymerisation disorder; mitosis initiation; cell proliferation;	
KW	cell growth; cell shape; cell rigidity; cell motility; DNA replication;	
KW	tumorigenesis; tumour survival; metastasis; ss.	
OS	Homo sapiens.	
XX		
EH	Location/Qualifiers	
FT	1..1791	
FT	/*tag= a	
FT	/product= "FEZ1"	
XX		
PN	WO200050565-A2.	
XX		
PD	31-AUG-2000.	
XX		
PF	25-FEB-2000; 2000WO-US04950.	
XX		
PR	25-FEB-1999; 99US-0121537.	
XX		
PA	(UYJE-) UNIV JEFFERSON THOMAS.	
XX		
PI	Croce CM, Ishii H;	
XX		
DR	WPI; 2000-558396/51.	
XX	P-PSDB; AAB08715.	
PT	New polynucleotide homologous with a portion of one strand of the human	
PT	FEZ1 gene, useful for alleviating abnormal cell proliferation such as	
XX	cancer -	
PS	Claim 7; Fig 5I; 255pp; English.	
XX		
CC	The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a	
CC	tumour suppressor gene, located at chromosome location 8p22. Decreased	
CC	or no expression of FEZ1 is detected in a variety of cancer cells.	
CC	Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1	
CC	also interacts with tubulin, with microtubules, and with protein	
CC	EF1-gamma. Post-translational phosphorylation and dephosphorylation	
CC	modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene	
CC	expression are useful for inducing cells to proliferate. Compounds	
CC	which modulate FEZ1 association with tubulin are useful for alleviating	
CC	tubulin hyper- or hypo- polymerisation disorders, such as those	
CC	associated with aberrant initiation of mitosis, modulation of the	
CC	initiation and rate of cell proliferation and cell growth, modulation of	
CC	cell shape, cell rigidity, cell motility, rate and stage of cellular	
CC	DNA replication, intracellular distribution of organelles, metastatic	
CC	potential of cell and cellular transformation from a non-cancerous to	
CC	cancerous phenotype. Compounds which modulate FEZ1 binding and	
CC	phosphorylation are also useful for alleviating a disorder, such as	
CC	tumorigenesis, tumour survival, growth and metastasis.	
XX	Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;	
SQ		



Db 877 ATCTCCAGCAGCTGCGAGCATCCAGGAGCTGGAGCAGAAAGCTCTTGGAGAGGGGCG 936  
Qy 421 GCCCTCCAGAGCTGCGAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGCGCTAC 480  
Db 937 GCCCTCCAGAGCTGCGAGCTTTGAGGAGAGGAGCTTGCTCCAGCTGCGCTAC 996  
Qy 481 GAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGGAGCCCAAGGGGCGCAAC 540  
Db 997 GAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGGAGCCCAAGGGGCGGAC 1056  
Qy 541 AAGTCAAGAGGCTGCGAGAGAGCGCGCGCGAGGCTGCTGCACTGCAAGTA 600  
Db 1057 AAGTCAAGAGGCTGCGAGAGAGCGCGCGCGAGGCTGCTGCACTGCAAGTA 1116  
Qy 601 CTGAGCTTACAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAG 660  
Db 1117 CTGAGCTTACAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAG 1176  
Qy 661 CAGGAGCTTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACGAGCTTCGGCCCC 720  
Db 1177 CAGGAGCTTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACGAGCTTCGGCCCC 1236  
Qy 721 CGCTGGAGGAGACCCAGTGGAGGTGAGG 751  
Db 1237 CGCTGGAGGAGACCCAGTGGAGGTGAGG 1267

## RESULT 10

AAS30595  
ID AAS30595 standard; cDNA; 560 BP.

XX AAS30595;

XX 21-NOV-2001 (first entry)

DE DNA encoding novel lung cancer antigen, Seq ID No 17.

XX Human; lung cancer; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virucide; fungicide;  
KW ophthalmological; autoimmune disease; rheumatoid arthritis; neoplasm;  
KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; infection; skin aging;  
KW ocular disorder; wound healing; organ transplantation; ss.

XX Homo sapiens.

XX WO200155300-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01238.

XX 31-JAN-2000; 2000US-0179065.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465565/50.

XX P-PSDB; AAU18935.

XX Isolated nucleic acid molecule encoding a lung cancer antigen is used  
PT in preventing, treating or ameliorating a medical condition -

XX Claim 1; SEQ ID No 17; 475pp; English.

XX The invention relates to novel isolated lung cancer antigen  
CC polynucleotides (I) and polypeptides (II). (I) and (II) are used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are  
CC also used in diagnosing a pathological condition or susceptibility to a  
CC pathological condition, in particular, lung cancer. The antibodies to

CC (II) can also be used in alleviating symptoms associated with the  
CC disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme  
CC linked immunosorbent assays (ELISA). Disorders which are diagnosed or  
CC treated include autoimmune diseases e.g. rheumatoid arthritis,  
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,  
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders  
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.  
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi and  
CC ocular disorders e.g. corneal infection. The polypeptides can also be  
CC used to aid wound healing and epithelial cell proliferation, to prevent  
CC skin aging due to sunburn, to maintain organs before transplantation, for  
CC supporting cell culture of primary tissues, to regenerate tissues and in  
CC chemotaxis. AAS30595-AAS30685 represent novel human lung cancer antigen  
CC coding sequences, PCR primers and related sequences of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at: ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 560 BP; 120 A; 196 C; 161 G; 83 T; 0 other;

Query Match 57.68; Score 463.4; DB 22; Length 560;  
Best Local Similarity 99.8%; Pred. No. 1.3e-78;  
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTCTGCACTCTCCCGGAGAGTGCCAGCCAGCTGCACCCCGCCCTCCAGAC 60

Db 96 GCCATCTCTGCACTCTCCCGGAGAGTGCCAGCCAGCTGCACCCCGCCCTCCAGAC 155

Qy 61 AAGCCCAAGGAGGAGGAGCTGAAGCCTTGCTCTGGGGCGCTGTGAGACTCCGGC 120

Db 156 AAGCCCAAGGAGGAGGAGCTGAAGCCTTGCTCTGGGGCGCTGTGAGACTCCGGC 215

Qy 121 CGGAACTCCATGTCCAGCTGCCACACAGCAGCAGCAGCAGCTACCGTGGACCCG 180

Db 216 CGGAACTCCATGTCCAGCTGCCACACAGCAGCAGCAGCAGCTACCGTGGACCCG 275

Qy 181 CTGTGCACACCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACACATCACCAG 240

Db 276 CTGTGCACACCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACACATCACCAG 335

Qy 241 GGCACTCTCTCCAGGACAGCAATGATGAGCTGAAGGCTGTGCTTCTCCGAGGA 300

Db 336 GGCACTCTCTCCAGGACAGCAATGATGAGCTGAAGGCTGTGCTTCTCCGAGGA 395

Qy 301 GGTAGCAAGCTGGGCGCACTCGAACAGGACAGCAAGGGCCCTCGTGTGCTCCGCTCCCC 360

Db 396 GGTAGCAAGCTGGGCGCACTCGAACAGGACAGCAAGGGCCCTCGTGTGCTCCGCTCCCC 455

Qy 361 ATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAACAGAGCTGTGAGAGGAGGCG 420

Db 456 ATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAACAGAGCTGTGAGAGGAGGCG 515

Qy 421 GCCCTCCAGAGCTGCGAGCGAGCTTTGAGGAGAGGAGCTTGCC 465

Db 516 GCCCTCCAGAGCTGCGAGCGAGCTTTGAGGAGAGGAGCTTGCC 560

## RESULT 11

AAS28127  
ID AAS28127 standard; cDNA; 560 BP.

XX AAS28127;

XX 07-NOV-2001 (first entry).

XX Novel cDNA encoding for human respiratory antigen #259.

XX Human; respiratory antigen; respiratory disorder; throat disorder;  
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;  
KW anti-allergic; anti-asthmatic; anti-inflammatory; olfactory;  
KW respiratory active; ss.  
XX Homo sapiens.

XX WO200155448-A1.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01333.  
XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234224.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249224.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249254.  
PR 17-NOV-2000; 2000US-0249255.  
PR 17-NOV-2000; 2000US-0249256.  
PR 17-NOV-2000; 2000US-0249257.  
PR 17-NOV-2000; 2000US-0249258.  
PR 17-NOV-2000; 2000US-0249259.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI INC.  
XX FA  
XX XX

PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-476224/51.  
 DR P-PSDB; AAU17943.  
 XX  
 PT Isolated polypeptide for treating, preventing and/or prognosing  
 PT disorders related to the respiratory system including respiratory  
 PT cancers and also for testing and detection e.g. diagnosis -  
 XX  
 PS Claim 4; SED ID No 269; 546pp; English.  
 XX  
 CC The present invention relates to the isolation of novel human  
 CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic  
 CC sequences encoding for these polypeptides. The sequences of the  
 CC invention are useful for preventing, treating and/or prognosing  
 CC disorders related to the respiratory system including throat  
 CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),  
 CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,  
 CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of  
 CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences  
 CC of the invention are useful in gene therapy and antisense therapy.  
 CC AAS27869-AAS28159 encode for novel human respiratory antigens.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 560 BP; 120 A; 196 C; 161 G; 83 T; 0 other;  
 Query Match 57.6%; Score 463.4; DB 22; Length 560;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-78;  
 Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCCATCTGCACTCTCCCGGAGAGTGCAGCACCAGCTGCACCCCGCCCTCCAGAC 60  
 Db 96 GCCATCTGCACTCTCCCGGAGAGTGCAGCACCAGCTGCACCCCGCCCTCCAGAC 155  
 QY 61 AAGCCCAAGGAGCAGAGCTGAAGCTGGCTGTCTGGGGCGCTGTAGACTCCGGC 120  
 Db 156 AAGCCCAAGGAGCAGAGCTGAAGCTGGCTGTCTGGGGCGCTGTAGACTCCGGC 215  
 QY 121 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCTACCGCTGAGACCG 180  
 Db 216 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCTACCGCTGAGACCG 275  
 QY 181 CTGGTACACCGCTGGGACCCACAGCGCTTTTGGGGGCTCGGCCACAAATACCCAG 240  
 Db 276 CTGGTACACCGCTGGGACCCACAGCGCTTTTGGGGGCTCGGCCACAAATACCCAG 335  
 QY 241 GCATCGTCTCCAGCAGCAGCAATGATGAGCTTGAAGGCTCTGTCTTCTCCGACGA 300  
 Db 336 GCATCGTCTCCAGCAGCAGCAATGATGAGCTTGAAGGCTCTGTCTTCTCCGACGA 395  
 QY 301 GGTAGCAAGCTGGGCACTCGAACAAGGAGCAGCAAGGGCCCTCTGTGTGCTCCGCC 360  
 Db 396 GGTAGCAAGCTGGGCACTCGAACAAGGAGCAGCAAGGGCCCTCTGTGTGCTCCGCC 455  
 QY 361 ATCTCCAGGAGCAGTGCAGCATCCAGAGCTGGACAGAGCTTTGGAGAGGAGGCG 420  
 Db 456 ATCTCCAGGAGCAGTGCAGCATCCAGAGCTGGAGAGCTTTGGAGAGGAGGCG 515  
 QY 421 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGGAGCTTGCC 465  
 Db 516 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGGAGCTTGCC 560  
 RESULT 12  
 ID AAA64511  
 XX AAA64511 standard; cDNA; 633 BP.  
 AC AAA64511;  
 XX  
 DT 02-JAN-2001 (first entry)  
 XX

DE Nucleotide sequence of truncated FEZ1 transcript E264162.  
 XX  
 KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
 KW tumour proliferation; tubulin; microtubule; protein E1-gamma;  
 KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
 KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
 KW tumorigenesis; tumour survival; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..633  
 FT /\*tag= a  
 FT /product= "truncated FEZ1"  
 XX  
 FN MO2000050565-A2.  
 XX  
 PD 31-AUG-2000.  
 XX  
 PF 25-FEB-2000; 2000WO-US04950.  
 XX  
 PR 25-FEB-1999; 99US-0121537.  
 XX  
 PA (UJJE-) UNIV JEFFERSON THOMAS.  
 XX  
 PI Croce CM, Ishii H;  
 XX  
 DR WPI; 2000-558396/51.  
 DR P-PSDB; AAB08718.  
 XX  
 PT New polynucleotide homologous with a portion of one strand of the human  
 PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
 PT cancer -  
 XX  
 PS Disclosure; Fig 5D; 255pp; English.  
 XX  
 CC The present sequence encodes a truncated human FEZ1 polypeptide. The  
 CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
 CC suppressor gene, located at chromosome location 8p22. Decreased  
 CC or no expression of FEZ1 is detected in a variety of cancer cells.  
 CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
 CC also interacts with tubulin, with microtubules, and with protein  
 CC E1-gamma. Post-translational phosphorylation and dephosphorylation  
 CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
 CC expression are useful for inducing cells to proliferate. Compounds  
 CC which modulate FEZ1 association with tubulin are useful for alleviating  
 CC tubulin hyper- or hypo- polymerisation disorders, such as those  
 CC associated with aberrant initiation of mitosis, modulation of the  
 CC initiation and rate of cell proliferation and cell growth, modulation of  
 CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
 CC DNA replication, intracellular distribution of organelles, metastatic  
 CC potential of cell and cellular transformation from a non-cancerous to  
 CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
 CC phosphorylation are also useful for alleviating a disorder, such as  
 CC tumorigenesis, tumour survival, growth and metastasis.  
 XX  
 SQ Sequence 633 BP; 136 A; 217 C; 175 G; 105 T; 0 other;  
 Query Match 19.6%; Score 157.6; DB 21; Length 633;  
 Best Local Similarity 92.2%; Pred. No. 6e-21;  
 Matches 166; Conservative 0; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 GCCATCTGCACTCTCCCGGAGAGTGCAGCACCAGCTGCACCCCGCCCTCCAGAC 60  
 Db 406 GCCATCTGCACTCTCCCGGAGAGTGCAGCACCAGCTGCACCCCGCCCTCCAGAC 465  
 QY 61 AAGCCCAAGGAGCAGAGCTGAAGCTGGCTGTCTGGGGCGCTGTAGACTCCGGC 120  
 Db 466 AAGCCCAAGGAGCAGAGCTGAAGCTGGCTGTCTGGGGCGCTGTAGACTCCGGC 525  
 QY 121 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCTACAGTGCAGTGCACCG 180  
 Db 526 CGGAACCTCCATGCTCCAGCTGCCACACACAGCAGCAGCTTCGAGGTTGACCTG 585

RESULT 13  
ABK62287/c  
ID ABK62287 standard; cDNA; 561 BP.  
XX  
AC ABK62287;  
XX  
AT 18-JUN-2002 (first entry)  
XX  
DE Rat sequence differentially expressed in response to a hepatotoxin #194.  
XX  
KW Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;  
XX differential expression; centrilobular necrosis; steatosis.  
XX  
KW Rattus norvegicus.  
OS  
XX  
PN WO200210453-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 30-JUL-2001; 2001WO-US23872.  
XX  
PR 31-JUL-2000; 2000US-222040P.  
PR 02-NOV-2000; 2000US-244880P.  
PR 11-MAY-2001; 2001US-290029P.  
PR 15-MAY-2001; 2001US-290645P.  
PR 22-MAY-2001; 2001US-292336P.  
PR 06-JUN-2001; 2001US-295798P.  
PR 13-JUN-2001; 2001US-297457P.  
PR 19-JUN-2001; 2001US-298884P.  
PR 09-JUL-2001; 2001US-303459P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Mendrick D, Porter MW, Johnson KR, Castle AL, Elashoff MR;  
XX  
DR WPI; 2002-241625/29.  
XX  
PT Predicting toxic effects of compounds or the progression of these toxic  
PT effects by determining the changes in gene expression in tissues or  
PT cells exposed to the toxin and comparing these to gene expression in  
PT unexposed tissues or cells -  
XX  
PS Claim 1; Seq ID No 194; 239pp; English.  
XX  
CC The invention relates to methods for predicting toxic effects of  
CC compounds or the progression of these toxic effects by determining the  
CC global changes in gene expression in tissues or cells exposed to the  
CC toxin and comparing these to gene expression in unexposed tissues or  
CC cells. Also included are methods of predicting at least one toxic  
CC effect of a compound or progression of a toxic effect, preferably the  
CC hepatotoxicity of a compound, comprising detecting the level of  
CC expression in a tissue or cell sample exposed to the compound of two or  
CC more genes listed in the specification, where differential expression of  
CC the genes is indicative of at least one toxic effect or progression.  
CC The method can also be used to identify an agent which modulates the  
CC toxic response and predict cellular pathways that a compound modulates  
CC in a cell. The methods utilise a set of at least two probes (on a solid  
CC support in kit form), where each of the probes comprises a sequence that  
CC specifically hybridises to a gene listed in the specification, a computer  
CC system comprising a database containing information identifying the  
CC expression level in a tissue or cell sample exposed to a hepatotoxin of a  
CC set of genes comprising at least two genes listed in the specification,  
CC and a user interface to view the information used to present information  
CC identifying the expression level in a tissue or cell of at least one gene  
CC listed in the specification. The method is useful for elucidating global  
CC changes in gene expression and for identifying toxicity markers in  
CC tissues or cell exposed to a known toxin. The genes may be used as  
CC toxicity markers in drug screening and toxicity assays. The genes and  
CC gene expression information may be used as diagnostic markers for the  
CC prediction or identification of the physiological state of tissue or cell  
CC sample that has been exposed to a compound or agent. Hepatotoxicity

is characterised by centrilobular necrosis and steatosis. The present  
sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
which is differentially expressed in response to a hepatotoxic agent.  
Sequence 561 BP; 151 A; 143 C; 127 G; 140 T; 0 other;  
Query Match 14.0%; Score 112.4; DB 24; Length 561;  
Best Local Similarity 77.0%; Pred. No. 2e-12;  
Matches 137; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
QY 367 ACGGACGAGTCAGATCCAGAGCTGGAACAGAAAGCTGTTGGAGAGGAGGGCCCTC 426  
DB 561 ACGGATGAGTCACCAATTCAGAGCTGGAGCAGAAGCTGCTGCAGCGGAGACTGCACTA 502  
QY 427 CAGAAGCTGCAGCGCAGCTTTGAGGAGAAGAGCTTGCCTCCAGCTGCGCTTACGAGGAG 486  
DB 501 CAGAAGCTTACAGCGCAGCTTTCATGAGNAGAGTTTGCCTCTGGCCACACCTTCGAGGAG 442  
QY 487 CGGCGGCGGCTGCAGGACGAGCTGGAGGCGGCGGAGCCCAAGGCGGCAACAGC 544  
DB 441 CGGCCACGGCGGACTAGAGACGAAGCTGGAGTGCTGGAACCTAGAGCAAGCTGAAGC 384  
RESULT 14  
AAV55830  
ID AAV55830 standard; DNA; 795 BP.  
XX  
AC AAV55830;  
XX  
DT 18-NOV-1998 (first entry)  
XX  
DE FLGA insert stabilising polypeptide encoding DNA.  
XX  
KW Fusion protein; stabilising polypeptide; proteolytic degradation;  
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;  
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;  
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;  
KW cancer; pathological condition; ss.  
XX  
OS Epstein-barr virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..788  
FT /tag= a  
FT /product= "stabilising polypeptide"  
XX  
PN WO9822577-A1.  
XX  
PD 28-MAY-1998.  
XX  
PF 17-NOV-1997; 97WO-IB01508.  
XX  
PR 25-JUN-1997; 97US-0048945.  
PR 15-NOV-1996; 96US-0030986.  
XX  
PA (MASU/) MASUCCI M G.  
XX  
PI Masucci MG;  
XX  
DR WPI; 1998-312463/27.  
XX P-PSDB; AAW79128.  
XX  
PT New fusion proteins resistant to proteolytic degradation -  
PT comprising a core protein with a stabilising polypeptide comprising  
XX a peptide sequence containing glycine repeats  
PS Disclosure; Fig 3; 120pp; English.  
XX  
CC This DNA encodes a stabilising polypeptide and is the FLGA insert of the  
CC invention. The invention provides a method for increasing the resistance  
CC of a core protein to proteolytic degradation that comprises linking or  
CC inserting onto or into the core protein a stabilising polypeptide of  
CC formula [(Glya)x(Glyb)y(Glyc)z]n where Glya, Glyb, Glyc are 1-6

CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,  
CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not  
CC be identical from n repeat to n repeat. Alternatively a nucleic acid  
CC encoding the stabilising polypeptide can be linked onto or inserted into  
CC a nucleic acid encoding a core protein. The fusion proteins of the  
CC invention are more resistant to degradation by proteases and, thus, have  
CC a longer half-life than the unfused core protein. The products can be  
CC used for treating autoimmune diseases, cancer and inflammation. In  
CC particular, the core protein may be an IkappaB regulator protein for the  
CC treatment of inflammatory bowel disease, or a nitroreductase protein  
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer  
CC or other pathological conditions. The fusion proteins can also be used in  
CC diagnostic methods such as in vivo imaging.

XX Sequence 795 BP; 200 A; 104 C; 478 G; 13 T; 0 other;

Query Match 10.3%; Score 82.8; DB 19; Length 795;  
Best Local Similarity 51.6%; Pred. No. 7.7e-07;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
QY 386 AGGAGCTGGAACAGAACTGTGTGGAGAGGGGCGCCCTCCAGAACTGTCAGCGCAGCT 445  
DB 274 AGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 333  
QY 446 TTGAGGAGAGGAGCTTCCCTCCAGCTGGCTACGAGGAGGCGCGCGCGCTGCAGGG 505  
DB 334 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 393  
QY 506 ACAGCTGGAGGGCGCGGAGGCCCAAGGCGGCAACAGCTCAAGCAGGCGCTCGCAGAGA 565  
DB 394 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 453  
QY 566 GCCAGCGCGGCGAGCAGCTCTGCACTGCGAGGTACTGAGTTCAGCTTCAGCAGAGAGCGCGC 625  
DB 454 AGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 513  
QY 626 AGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAGCAGAGCTGCTGAGACCAAGCTCA 685  
DB 514 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 573  
QY 686 GGTCTTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGGAGACCCAGTGGGAGG 745  
DB 574 AGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 633  
QY 746 TGAGGC 751  
DB 634 AGGGGC 639

## RESULT 15

AAV55831  
ID AAV55831 standard; DNA; 799 BP.

XX  
AC AAV55831;

XX  
DT 18-NOV-1998 (first entry)

XX Nucleotide sequence of the stabilising sequence-encoding insert.

XX Fusion protein; stabilising polypeptide; proteolytic degradation;  
KW resistance; half-life; autoimmune disease; inflammation; nitro drug;  
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;  
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;  
KW cancer; pathological condition; ss.

XX Epstein-Barr virus.

XX WO9822577-A1.

XX 28-MAY-1998.

XX 17-NOV-1997; 97WO-IB01508.

,XX

PR 25-JUN-1997; 97US-0048945.  
PR 15-NOV-1996; 96US-0030986.

XX (MASU//) MASUCCI M G.

XX Masucci MG;

XX WPI; 1998-312463/27.

XX New fusion proteins resistant to proteolytic degradation -  
PT comprising a core protein with a stabilising polypeptide comprising  
PT a peptide sequence containing glycine repeats

XX Disclosure; Fig 4B; 120pp; English.

CC This is a nucleotide sequence of the stabilising sequence-encoding  
CC insert. The invention provides a method for increasing the resistance  
CC of a core protein to proteolytic degradation that comprises linking or  
CC inserting onto or into the core protein a stabilising polypeptide of  
CC formula [(Glya)X(Glyb)Y(Glyc)Z]n where Glya, Glyb, Glyc are 1-6  
CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,  
CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not  
CC be identical from n repeat to n repeat. Alternatively a nucleic acid  
CC encoding the stabilising polypeptide can be linked onto or inserted into  
CC a nucleic acid encoding a core protein. The fusion proteins of the  
CC invention are more resistant to degradation by proteases and, thus, have  
CC a longer half-life than the unfused core protein. The products can be  
CC used for treating autoimmune diseases, cancer and inflammation. In  
CC particular, the core protein may be an IkappaB regulator protein for the  
CC treatment of inflammatory bowel disease, or a nitroreductase protein  
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer  
CC or other pathological conditions. The fusion proteins can also be used in  
CC diagnostic methods such as in vivo imaging.

XX Sequence 799 BP; 201 A; 106 C; 479 G; 13 T; 0 other;

Query Match 10.3%; Score 82.8; DB 19; Length 799;  
Best Local Similarity 51.6%; Pred. No. 7.8e-07;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 386 AGGAGCTGGAACAGAACTGTGTGGAGAGGGGCGCCCTCCAGAACTGTCAGCGCAGCT 445  
DB 277 AGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 336  
QY 446 TTGAGGAGAGGAGCTTCCCTCCAGCTGGCTACGAGGAGGCGCGCGCGCTGCAGGG 505  
DB 337 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 396  
QY 506 ACAGAGCTGGAGGGCGCGGAGGCCCAAGGCGGCAACAGCTCAAGCAGGCGCTCGCAGAGA 565  
DB 397 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 456  
QY 566 GCCAGCGCGGCGCAGCAGCTTCTGCACCTGCGAGGTACTGCGAGCTTCAGCAGAGAGCGCGC 625  
DB 457 AGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 516  
QY 626 AGCTCCGCGAGGAGCTCGAGAGCTCATGAGGAGCAGGAGCTCTGCTGAGAGCAGCTCA 685  
DB 517 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 576  
QY 686 GGTCTTACGAGAGGAGAGAACCAAGCTTCGGCCCGCGCTGAGAGGAGACCCAGTGGGAGG 745  
DB 577 AGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 636  
QY 746 TGAGGC 751  
DB 637 AGGGGC 642

Search completed: June 14, 2003, 23:43:04  
Job time : 188.106 secs





```
Qy 746 TGAGGC 751
Db 879 AGGGGC 884

RESULT 2
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2

Query Match 10.3%; Score 82.8; DB 3; Length 2580;
Best Local Similarity 51.6%; Pred. No. 3.1e-08;
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAACAGAAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCT 445
Db 902 AGGGCCAGGAGCAGGAGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961
Qy 446 TTGAGGAGAGGAGCTTCCCTCCAGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 505
Db 962 AGGAGGGGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1021
Qy 506 ACAGAGCTGGAGGGCCCGGAGCCCAAGCGGCAACAGCTCAAGCAGGCTCGCAGAGAAGA 565
Db 1022 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1081
Qy 566 GCCAGCGCGCAGCAGAGTCTGTCACCTGCAAGTATGAGGAGCAGGAGCTGCTGAGACCAAGCTCA 625
Db 1082 AGGGCCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1141
Qy 626 AGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGCAGGAGCTGCTGAGACCAAGCTCA 685
Db 1142 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1201
Qy 686 GGTCTTACGAGAGGGAGAGACCAAGCTTCGGGCGCGCTGGAGGAGACCCAGTGGGAGG 745

Db 1202 AGGGCAGGAGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1261
Qy 746 TGAGGC 751
Db 1262 AGGGGC 1267

RESULT 3
US-09-359-081-2
; Sequence 2, Application US/09359081
; Patent No. 6316223
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; TITLE OF INVENTION: System
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/359,081
; FILING DATE: 22-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/050,863
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-359-081-2

Query Match 10.3%; Score 82.8; DB 4; Length 2580;
Best Local Similarity 51.6%; Pred. No. 3.1e-08;
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

Qy 386 AGGAGCTGGAACAGAAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCT 445
Db 902 AGGGCCAGGAGCAGGAGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961
Qy 446 TTGAGGAGAGGAGCTTCCCTCCAGCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 505
Db 962 AGGAGGGGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1021
Qy 506 ACAGAGCTGGAGGGCCCGGAGCCCAAGCGGCAACAGCTCAAGCAGGCTCGCAGAGAAGA 565
Db 1022 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1081
Qy 566 GCCAGCGCGCAGCAGAGTCTGTCACCTGCAAGTATGAGGAGCAGGAGCTGCTGAGACCAAGCTCA 625
Db 1082 AGGGCCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1141
Qy 626 AGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGCAGGAGCTGCTGAGACCAAGCTCA 685
Db 1142 AGGAGCAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1201
Qy 686 GGTCTTACGAGAGGGAGAGACCAAGCTTCGGGCGCGCTGGAGGAGACCCAGTGGGAGG 745
```

Db 1082 AGGGCAGGAGCAGAGGGGGCAGGAGGAGGGGCGACAGGAGGAGGGGCGAGGGGCGAGGGGCG 1141  
Qy 626 AGCTCCGCGAGGAGCTCGAGAGCTTCATGAGGAGCAGGAGCTGCTGGAGACCAAGCTCA 685  
Db 1142 AGGAGCAGGAGGGCAGGAGGGGCGAGGAGGAGGGGCGAGGGGCGAGGAGGAGG 1201  
Qy 686 GGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGAGACCCAGTGGGAGG 745  
Db 1202 AGGGCAGGAGGGGCGAGGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGG 1261  
Qy 746 TGAGGC 751  
Db 1262 AGGGGC 1267

RESULT 4  
US-09-130-114-1/c  
; Sequence 1, Application US/09130114  
; Patent No. 5976807  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert A.  
; APPLICANT: Dama, Bassam B.  
; APPLICANT: Robbins, Alan K.  
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes  
; FILE REFERENCE: 0867/1D903US1  
; CURRENT APPLICATION NUMBER: US/09/130,114  
; CURRENT FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 5452  
; TYPE: DNA  
; ORGANISM: VEBNA  
US-09-130-114-1

Query Match 10.3%; Score 82.8; DB 2; Length 5452;  
Best Local Similarity 51.6%; Pred. No. 3.4e-08;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
Qy 386 AGGAGCTGGACAGAGCTGTGGAGAGGGGCGCCCTCCAGNAGCTGCGAGCGCAGCT 445  
Db 1903 AGGGCAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCG 1844  
Qy 446 TTGAGGAGAGGAGCTTCCCTCCAGCTTGGCTACGAGGAGCGCGCGCGGCTGCGAGG 505  
Db 1843 AGGAGGGCAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1784  
Qy 506 ACAGCTGGAGGGCCCGAGGAGCCAAAGCGCGCAACAGCTCAAGAGGCGCTCGCAGAGA 565  
Db 1783 AGGAGCAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGG 1724  
Qy 566 GCCAGCGCGCAGCAGCTCTGCACTTGCAGTACTGCACTTTCAGCTTTCAGCAGGAGCGGC 625  
Db 1723 AGGGCAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGG 1664  
Qy 626 AGCTCCGCGAGGAGCTCGAGAGCTTCATGAGGAGCAGGACCTGCTGGAGACCAAGCTCA 685  
Db 1663 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1604  
Qy 686 GGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGAGACCCAGTGGGAGG 745  
Db 1603 AGGGCAGGAGGGGCGAGGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1544  
Qy 746 TGAGGC 751  
Db 1543 AGGGGC 1538

RESULT 5  
US-08-910-647-1  
; Sequence 1, Application US/08910647  
; Patent No. 6251433

GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.  
TITLE OF INVENTION: Compositions and Methods for  
TITLE OF INVENTION: Polynucleotide Delivery  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/910,647  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-910-647-1

Query Match 10.3%; Score 82.8; DB 4; Length 9600;  
Best Local Similarity 51.6%; Pred. No. 3.7e-08;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
Qy 386 AGGAGCTGGAAACAGAGCTGTGGAGAGGGGAGGGGCGCCCTCCAGAGCTGAGCGCAGCT 445  
Db 948 AGGGCAGGAGCAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGG 1007  
Qy 446 TTGAGGAGAGGAGCTTGCCTCCAGCTTGGCTTACGAGGAGCGCGCGGCTGCGAGG 505  
Db 1008 AGGAGGGGCGAGGAGCAGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1067  
Qy 506 ACAGCTGGAGGGGCGCGAGGAGCCAAAGCGCGCAACAGCTCAAGCAGGCGCTCGCAGAGA 565  
Db 1068 AGGAGCAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGAGG 1127  
Qy 566 GCCAGCGCGCAGCAGGCTCTGCACTTGCAGTACTGCACTTTCAGCAGGAGAGCGGC 625  
Db 1128 AGGGCAGGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGG 1187  
Qy 626 AGCTCCGCGAGGAGCTCGAGAGCTTCATGAGGAGCAGGACCTGCTGGAGACCAAGCTCA 685  
Db 1188 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGG 1247  
Qy 686 GGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGAGAGAGACCCAGTGGGAGG 745  
Db 1248 AGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCG 1307  
Qy 746 TGAGGC 751  
Db 1308 AGGGGC 1313

RESULT 6  
US-09-620-925-1  
; Sequence 1, Application US/09620925

Patent No. 6468986  
GENERAL INFORMATION:  
APPLICANT: Zuckermann et al.  
TITLE OF INVENTION: Compositions and Methods for Polynucleotide Delivery  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Chiron Corporation  
STREET: 4560 Horton Street  
CITY: Emeryville  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 94608-2916  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION NUMBER: US/09/620,925  
FILING DATE: 21-Jul-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/910,647  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (Genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-620-925-1

Query Match 10.3%; Score 82.8; DB 4; Length 9600;  
Best Local Similarity 51.6%; Pred. No. 3.7e-08;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
Qy 386 AGGAGCTGGACAGAACTGTGTGGAGAGGGGCGCCCTCCAGAGCTGCAGCGCAGCT 445  
Db 948 AGGGCAGGAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1007  
Qy 446 TTGAGGAGAGGAGCTTCCCTCCAGCTTGGAGAGGGGCGCCCTCCAGAGCTGCAGGG 505  
Db 1008 AGGAGGGCAGGAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1067  
Qy 506 ACAGCTGGAGGGCGGAGCGCAAGCGCGCAAGCTCAAGCAGCGCTCGCAGAGA 565  
Db 1068 AGGAGCAGAGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1127  
Qy 566 GCCAGCGCGCAGCAGGCTTCTGCACTTCCAGGTTCTGAGGTTCTGAGGAGGAGGAGG 625  
Db 1128 AGGGCAGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1187  
Qy 626 AGCTCCGCGCAGGAGCTCAGAGCTTCATGAGGAGCAGGAGCTGCTGAGACCAAGCTCA 685  
Db 1188 AGGAGCAGAGGGGCGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1247  
Qy 686 GGTCTCAGAGGAGGAGAACCAAGCTTCCGCGCGCGCTGAGGAGGAGGAGGAGGAGG 745  
Db 1248 AGGGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1307  
Qy 746 TGAGGC 751  
Db 1308 AGGGGC 1313

## RESULT 7

US-07-884-811-15  
Sequence 15, Application US/07884811  
Patent No. 5316921  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.  
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/884,811  
FILING DATE: 19920518  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 755.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10596 bases  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-884-811-15

Query Match 10.3%; Score 82.8; DB 1; Length 10596;  
Best Local Similarity 51.6%; Pred. No. 3.8e-08;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
Qy 386 AGGAGCTGGACAGAACTGTGTGGAGAGGGGCGCCCTCCAGAGCTGCAGCGCAGCT 445  
Db 2482 AGGGCAGGAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2541  
Qy 446 TTGAGGAGAGGAGCTTCCCTCCAGCTTGGAGAGGGGCGCCCTCCAGAGCTGCAGGG 505  
Db 2542 AGGAGGGCAGGAGCAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2601  
Qy 506 ACAGCTTGGAGGGGCGGAGAGCCCAAGCGCGCAACAAGCTCAAGCAGGCTTCGAGAGA 565  
Db 2602 AGGAGCAGGAGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2661  
Qy 566 GCCAGCGCGCAGCAGGCTTCCAGCTTGCAGCTTGCAGCTTGCAGCTTGCAGGAGCGGC 625  
Db 2662 AGGGCAGGAGGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2721  
Qy 626 AGCTCCGCGCAGGAGCTCAGAGCTTCATGAGGAGCAGGAGCTGCTGAGAGCAAGCTCA 685  
Db 2722 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2781  
Qy 686 GGTCTTACGAGAGGAGAGAACCAAGCTTTCGGCCCGCGCTGAGGAGGAGGAGGAGGAGG 745  
Db 2782 AGGGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2841

Qy 746 TGAGGC 751  
Db 2842 AGGGGC 2847

## RESULT 8

US-07-885-971-15  
; Sequence 15, Application US/07885971  
; Patent No. 5328837  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.  
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/885,971  
; FILING DATE: 19920518  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dregler, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: 779  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10596 bases  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-07-885-971-15

Query Match 10.3%; Score 82.8; DB 1; Length 10596;  
Best Local Similarity 51.6%; Pred. No. 3.8e-08;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
Qy 386 AGGAGCTGGAACAGAGCTGTTGGAGAGGGGCGCCCTCCAGAGCTGCAGCGCAGCT 445  
Db 2482 AGGGGACAGCAGGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2541  
Qy 446 TTGAGGAGAGAGCTTGCCTCCAGCTGGCTTACGAGAGCGGCGCGCGCTGCAGGG 505  
Db 2542 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2601  
Qy 506 ACAGCTGGAGGGCCCGAGGCGCAAGAGCGGCAACAGCTCAAGAGCGCTTCGAGAGA 565  
Db 2602 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2661  
Qy 566 GCCAGCGGCGCAGCAGCTCTGCACCTGCAGGTACTGCAGCTTCAGCAGGAGAGCGGC 625  
Db 2662 AGGGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2721  
Qy 626 AGCTCCGACAGGAGCTTCAGAGCGCTTCATGAAAGGAGCAGGACCTGCTGAGACCAAGCTCA 685  
Db 2722 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2781  
Qy 686 GGTCTTACGAGAGGAGAGAACAGCTTTCGCGCCCGCGCTGAGAGGAGACCAGTGGGAGG 745

Db 2782 AGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2841  
Qy 746 TGAGGC 751  
Db 2842 AGGGGC 2847

## RESULT 9

US-08-087-783A-15  
; Sequence 15, Application US/08087783A  
; Patent No. 5547856  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.  
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/087,783A  
; FILING DATE: 13-Jul-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/884811  
; FILING DATE: 18-MAY-92  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/885971  
; FILING DATE: 18-MAY-92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Marschang, Diane L.  
; REGISTRATION NUMBER: 35,600  
; REFERENCE/DOCKET NUMBER: P0755779P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-5416  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10596 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
US-08-087-783A-15

Query Match 10.3%; Score 82.8; DB 1; Length 10596;  
Best Local Similarity 51.6%; Pred. No. 3.8e-08;  
Matches 189; Conservative 0; Mismatches 177; Indels 0; Gaps 0;  
Qy 386 AGGAGCTGGAACAGAGCTGTTGGAGAGGGGCGCCCTCCAGAGCTGCAGCGCAGCT 445  
Db 2482 AGGGGACAGCAGGAGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2541  
Qy 446 TTGAGGAGAGAGAGCTTGCCTCCAGCTGGCTTACGAGAGCGGCGCGCGCTGCAGGG 505  
Db 2542 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2601  
Qy 506 ACAGCTGGAGGGCCCGAGGCGCAAGAGCGGCAACAGCTCAAGAGCGCTTCGAGAGA 565  
Db 2602 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2661  
Qy 566 GCCAGCGGCGCAGCAGCTTCGACCTGCAGGTACTGCAGCTTCAGCAGGAGAGCGGC 625  
Db 2662 AGGGGACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2721













GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 15, 2003, 02:18:56 ; Search time 117.039 Seconds  
(without alignments)  
9947.427 Million cell updates/sec

Title: US-09-513-888c-1\_COPY\_1707\_2510

Perfect score: 804

Sequence: 1 gccatctgactctctcccc.....tttggcgccagtaccocct 804

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1029858 seqs, 724030393 residues

Total number of hits satisfying chosen parameters: 2059716

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	802.4	99.8	2411	9	US-09-764-904-89
2	802.4	99.8	2411	9	US-10-091-548-89
3	802.4	99.8	2411	9	US-10-074-095-1133
4	802.4	99.8	2411	10	US-09-764-860-1133
5	463.4	57.6	560	9	US-09-764-904-17
6	463.4	57.6	560	9	US-10-091-548-17
7	463.4	57.6	560	9	US-10-074-095-269
8	463.4	57.6	560	10	US-09-764-860-269
9	151	18.8	5257	9	US-10-171-581-63
10	112.4	14.0	561	10	US-09-917-800A-194
11	67	8.3	1617	10	US-09-925-300-60
12	67	8.3	2108	10	US-09-962-832-225
13	57.8	7.2	8459	9	US-09-563-728A-31
14	57.8	7.2	8459	10	US-09-817-913-8
15	57.8	7.2	8459	10	US-09-817-538-8
16	55.6	6.9	14800	10	US-09-954-456-1601
17	55.4	6.9	4366	12	US-10-044-090-668
18	55	6.8	596	10	US-09-864-761-8648
19	55	6.8	2145	9	US-10-153-668-125

20	55	6.8	2145	9	US-10-153-668-127	Sequence 127, App
21	54.4	6.8	2313	9	US-09-854-133-157	Sequence 157, App
22	54.4	6.8	2313	10	US-09-738-973-157	Sequence 157, App
23	54.4	6.8	2314	9	US-09-764-868-12	Sequence 12, Appl
24	54.4	6.8	2314	10	US-09-778-927A-16	Sequence 16, Appl
25	53.8	6.7	4282	10	US-09-864-761-25366	Sequence 25366, A
26	53.8	6.7	2785	10	US-09-925-302-122	Sequence 122, App
27	53.6	6.7	6604	10	US-09-880-107-1748	Sequence 1748, App
28	53.4	6.6	1662	9	US-09-292-862-1	Sequence 1, Appli
29	53.2	6.6	1649	9	US-09-789-390-62	Sequence 62, Appl
30	53.2	6.6	1787	9	US-09-789-390-66	Sequence 66, Appl
31	53.2	6.6	3822	9	US-09-789-390-10	Sequence 10, Appl
32	53.2	6.6	3971	9	US-10-270-333-59	Sequence 59, Appl
33	53.2	6.6	3999	9	US-09-789-390-6	Sequence 6, Appli
34	53.2	6.6	4133	9	US-09-291-417-11	Sequence 11, Appl
35	53.2	6.6	7503	9	US-10-270-333-58	Sequence 58, Appl
36	53	6.6	422	9	US-09-854-133-337	Sequence 337, App
37	53	6.6	422	10	US-09-738-973-337	Sequence 337, App
38	52.8	6.6	22585	9	US-09-764-891-6987	Sequence 6987, Ap
39	52.6	6.5	2614	9	US-09-822-846-491	Sequence 491, App
40	52.6	6.5	4033	9	US-10-029-115-5	Sequence 5, Appli
41	52.2	6.5	1852	10	US-09-969-852-4	Sequence 4, Appli
42	51.2	6.4	1827	9	US-09-974-298-174	Sequence 174, App
43	51.2	6.4	1827	10	US-09-919-172-48	Sequence 48, Appl
44	51.2	6.4	6457	10	US-09-880-107-3389	Sequence 3389, Ap
45	51	6.3	520	9	US-10-184-644-332	Sequence 332, App

#### ALIGNMENTS

##### RESULT 1

US-09-764-904-89  
; Sequence 89, Application US/09764904  
; Patent No. US20020173454A1

GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA122

; CURRENT APPLICATION NUMBER: US/09/764, 904

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 89

; LENGTH: 2411

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-904-89

Query Match 99.8%; Score 802.4; DB 9; Length 2411;  
Best Local Similarity 99.9%; Pred. No. 6.8e-187;  
Matches 803; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	GCCATCTGCACTCTCTCCCGGAGAGTGCCACGACGCTGCACCCCGCCCTCCAGAC	60
Db	1367	GCCATCTGCACTCTCTCCCGGAGAGTGCCACGACGCTGCACCCCGCCCTCCAGAC	1426
Qy	61	AAGCCCAAGGAGGAGGAGCTGAAGCCTGGCTGTGTCTTGGGGGCTGTAGACTCCGGC	120
Db	1427	AAGCCCAAGGAGGAGGAGCTGAAGCCTGGCTGTGTCTTGGGGGCTGTAGACTCCGGC	1486
Qy	121	CGGAATCTCCATGTCTCAGCTGCCACACACAGCAGCAGCAGCTACCACTGACCCG	180
Db	1487	CGGAATCTCCATGTCTCAGCTGCCACACACAGCAGCAGCAGCTACCACTGACCCG	1546
Qy	181	CTGTGTCAACCCGTTGGGACCCCAAGCCGTTTTGGGGCTCCGCCCAACATCACCAG	240
Db	1547	CTGTGTCAACCCGTTGGGACCCCAAGCCGTTTTGGGGCTCCGCCCAACATCACCAG	1606
Qy	241	GGCATCTCTTCCAGGACGACCAATATAGCTGAAGGCTCTGTCTTCTTCGACCGA	300
Db	1607	GGCATCTCTTCCAGGACGACCAATATAGCTGAAGGCTCTGTCTTCTTCGACCGA	1666

Qy	301	GGTAGCAAGCTGGGCCACTCGAAACAAGCGCAGACAAGGGCCCTCGTGTGTCCGCTCCCC	360
Db	1667	GGTAGCAAGCTGGGCCACTCGAAACAAGCGCAGACAAGGGCCCTCGTGTGTCCGCTCCCC	1726
Qy	361	ATCTCCACGACGAGTGCACATCCAGGACTCGAAACAGAGCTGTTGGAGAGGGAGGGC	420
Db	1727	ATCTCCACGACGAGTGCACATCCAGGACTCGAAACAGAGCTGTTGGAGAGGGAGGGC	1786
Qy	421	GCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCTCCAGCTGGCGCTAC	480
Db	1787	GCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCTCCAGCTGGCGCTAC	1846
Qy	481	GAGGAGCGCGCGCGCTCGAGGACGAGCTGGAGGGCCCGGAGCCCAAAGCGCGCAAC	540
Db	1847	GAGGAGCGCGCGCGCTCGAGGACGAGCTGGAGGGCCCGGAGCCCAAAGCGCGCAAC	1906
Qy	541	AAGCTCAAGCAGGCTCTCGCAGAAGAGCCAGCGCGCAGCAGGTCCTGCACTTCAGGTA	600
Db	1907	AAGCTCAAGCAGGCTCTCGCAGAAGAGCCAGCGCGCAGCAGGTCCTGCACTTCAGGTA	1966
Qy	601	CTGAGCTTCAGCAGGAGAACGCGCAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGAG	660
Db	1967	CTGAGCTTCAGCAGGAGAACGCGCAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGAG	2026
Qy	661	CAGGACCTGTGGAGCCAGCTCAGGTCCTACGAGGGGAGAGACCAAGTTCCGGCCCC	720
Db	2027	CAGGACCTGTGGAGCCAGCTCAGGTCCTACGAGGGGAGAGACCAAGTTCCGGCCCC	2086
Qy	721	CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGTTGGTGTGTCA	780
Db	2087	CGCTGGAGGAGACCCAGTGGAGGTGAGGCCACACAGGGCTCATGGTTGGTGTGTCA	2146
Qy	781	CGCGTTTGGCGCCAGTACCCCCCT	804
Db	2147	CGCGTTTGGCGCCAGTACCCCCCT	2170

## RESULT 2

```

US-10-091-548-89
; Sequence 89, Application US/10091548
; Publication No. US20030049703A1.
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA122C1
; CURRENT APPLICATION NUMBER: US/10/091,548
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 137
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-548-89

```

	Query Match.	99.8%	Score	802.4;	DB	9;	Length	2411;
	Best Local Similarity	99.9%;	Pred. No.	6.8e-187;				
	Matches	803;	Conservative	0;	Mismatches	1;	Indels	0;
								Gaps
Qy	1	GCCATCTGCAC	TCTCTCCCGAGAGTGC	CAGCACCAGCAGCTGC	CACCGGCCCTCC	GAC	60	
Db								
	1367	GCCATCTGCAC	TCTCTCCCGAGAGTGC	CAGCACCAGCAGCTGC	CACCGGCCCTCC	GAC	1426	
Qy	61	AAGCCCAAGAG	CAGCAGCTGA	AGCTTGGCTCTGTGGGCGCTGT	CAGACTCC	GGCG	120	
Db								
	1427	AAGCCCAAGAG	CAGCAGCTGA	AGCTTGGCTCTGTGGGCGCTGT	CAGACTCC	GGCG	1486	
Qy	121	CGGAATCTC	ATGTCCAGCTGCC	CACACAGCACCAGCAGCAGCTAC	CAGCTCGA	CCCG	180	
Db								
	1487	CGGAATCTC	ATGTCCAGCTGCC	CACACAGCACCAGCAGCAGCTAC	CAGCTCGA	CCCG	1546	

Qy	181	CTGGTCAACCCCGTGGGACCCCAAGCCGTTTGGGGGCTCCGCCCAACAATCATCCACG	240
Db	1547	CTGGTCAACCCCGTGGGACCCCAAGCCGTTTGGGGGCTCCGCCCAACAATCATCCACG	1606
Qy	241	GGCATCGTCTCCAGGACAGCAACATGATGAGCGCTGAAGGCTCTGTCTTCTCCGACGGA	300
Db	1607	GGCATCGTCTCCAGGACAGCAACATGATGAGCGCTGAAGGCTCTGTCTTCTCCGACGGA	1666
Qy	301	GGTAGCAAGCTGGGCCACTCGAAACAAGGCAGACAAGGGCCCTCGTGTGTCCGCTCCCCC	360
Db	1667	GGTAGCAAGCTGGGCCACTCGAAACAAGGCAGACAAGGGCCCTCGTGTGTCCGCTCCCCC	1726
Qy	361	ATCTCCACGACGAGTGCAGCATCTCCAGGAGCTGGAACAGAGTGTGTTGGAGAGGGAGGGC	420
Db	1727	ATCTCCACGACGAGTGCAGCATCTCCAGGAGCTGGAACAGAGTGTGTTGGAGAGGGAGGGC	1786
Qy	421	GCCTCCAGAAGCTGCAGCGCAGCTTTTGAGGAGAAGGAGCTTGCTCCAGCCTGGCCTAC	480
Db	1787	GCCTCCAGAAGCTGCAGCGCAGCTTTTGAGGAGAAGGAGCTTGCTCCAGCCTGGCCTAC	1846
Qy	481	GAGGAGCGCGCGCGCTTCAGGGACGAGCTGGAGGGCCCGGAGCCCAAGGGCGCAAC	540
Db	1847	GAGGAGCGCGCGCGCTTCAGGGACGAGCTGGAGGGCCCGGAGCCCAAGGGCGCAAC	1906
Qy	541	AAGCTCAAGCAGCGCTTCGCAAGAGAGCCAGCGCGCAGCAGGTCTCTGCACCTGCAGGTA	600
Db	1907	AAGCTCAAGCAGCGCTTCGCAAGAGAGCCAGCGCGCAGCAGGTCTCTGCACCTGCAGGTA	1966
Qy	601	CTGAGCTTCAGCAGGAGAAACGCGAGCTCCGCGCAGGAGCTCGAGAGCCTCATGAAGGAG	660
Db	1967	CTGAGCTTCAGCAGGAGAAACGCGAGCTCCGCGCAGGAGCTCGAGAGCCTCATGAAGGAG	2026
Qy	661	CAGGACCTGCTGAGACCAAGCTCAGTCTACGAGGGGAGAGACCGTTCGGGCCCC	720
Db	2027	CAGGACCTGCTGAGACCAAGCTCAGTCTACGAGGGGAGAGACCGTTCGGGCCCC	2086
Qy	721	CGCGTGGAGAGACCCAGTGGGAGGTGAGGCCACACAGGGCTCATGGTTGGTGGTCA	780
Db	2087	CGCGTGGAGAGACCCAGTGGGAGGTGAGGCCACACAGGGCTCATGGTTGGTGGTCA	2146
Qy	781	CGCGTTTGGCGCCAGTACCCCCCT	804
Db	2147	CGCGTTTGGCGCCAGTACCCCCCT	2170
RESULT 3			
US-10-074-095-1133			
; Sequence 1133, Application US/10074095			
; Publication No. US2003007704A1			
; GENERAL INFORMATION:			
; APPLICANT: Rosen et al.			
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies			
; FILE REFERENCE: PC008C1			
; CURRENT APPLICATION NUMBER: US/10/074,095			
; CURRENT FILING DATE: 2002-02-14			
; PRIOR APPLICATION NUMBER: 09/764,860			
; PRIOR FILING DATE: 2001-01-17			
; PRIOR APPLICATION NUMBER: 60/179,065			
; PRIOR FILING DATE: 2000-01-31			
; PRIOR APPLICATION NUMBER: 60/180,628			
; PRIOR FILING DATE: 2000-02-04			
; PRIOR APPLICATION NUMBER: 60/214,886			
; PRIOR FILING DATE: 2000-06-28			
; PRIOR APPLICATION NUMBER: 60/217,487			
; PRIOR FILING DATE: 2000-07-11			
; PRIOR APPLICATION NUMBER: 60/225,758			
; PRIOR FILING DATE: 2000-08-14			
; PRIOR APPLICATION NUMBER: 60/220,963			
; PRIOR FILING DATE: 2000-07-26			
; PRIOR APPLICATION NUMBER: 60/217,496			
; PRIOR FILING DATE: 2000-07-11			
; PRIOR APPLICATION NUMBER: 60/225,447			
; PRIOR FILING DATE: 2000-08-14			

### RESULT 3

US-10-074-095-1133  
Sequence 1133, Application US/10074095  
Publication No. US2003007704A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC008C1  
CURRENT APPLICATION NUMBER: US/10/074,095  
CURRENT FILING DATE: 2002-02-14  
PRIOR APPLICATION NUMBER: 09/764,860  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14

; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370

; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08



Db 1847 GAGGAGCGCGCGCGCTGAGGAGCGAGTGGAGGCGCGGCAAC 1906  
Qy 541 AAGCTCAAGCAGGCTCGCAGAGAGCAGCGCGCAGCAGGCTCTGCACCTGCAGTA 600  
Db 1907 AAGCTCAAGCAGGCTCGCAGAGAGCAGCGCGCGCAGCAGGCTCTGCACCTGCAGTA 1966  
Qy 601 CTGAGCTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGAG 660  
Db 1967 CTGAGCTTCAGCAGGAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGAG 2026  
Qy 661 CAGGACCTCTGGAGACCAAGCTCAGTCTTACGAGGAGAGACAGCTTCGGCCCC 720  
Db 2027 CAGGACCTCTGGAGACCAAGCTCAGTCTTACGAGGAGAGACAGCTTCGGCCCC 2086  
Qy 721 CGCTGGAGGAGACCCAGTGGAGGTGAGGTCACACAGCGCTCATGGTTGGGTGTC 780  
Db 2087 CGCTGGAGGAGACCCAGTGGAGGTGAGGTCACACAGCGCTCATGGTTGGGTGTC 2146  
Qy 781 GCGGTTTGGGCCCACTACCCCT 804  
Db 2147 GCGGTTTGGGCCCACTACCCCT 2170

## RESULT 5

US-09-764-904-17  
; Sequence 17, Application US/09764904  
; Patent No. US20020173454A1

; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA122

; CURRENT APPLICATION NUMBER: US/09/764,904

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 137

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 17

; LENGTH: 560

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-904-17

Query Match 57.6%; Score 463.4; DB 9; Length 560;  
Best Local Similarity 99.8%; Pred. No. 3.6e-104;  
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCTCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60  
Db 96 GCCATCTGCACTCTCTCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGAC 155  
Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 120  
Db 156 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 215  
Qy 121 CGGAATCCATGTCCAGCTCCACACACAGCAGCAGCAGCTACCACTGAGACCCG 180  
Db 216 CGGAATCCATGTCCAGCTCCACACACAGCAGCAGCAGCTACCACTGAGACCCG 275  
Qy 181 CTGCTACACCCGTGGAGCCACAGCGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 240  
Db 276 CTGCTACACCCGTGGAGCCACAGCGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 335  
Qy 241 GGCATCGTCTCCAGGACAGCAATGATGAGCTGTGAGGCTGTGCTCTCCGACGGA 300  
Db 336 GGCATCGTCTCCAGGACAGCAATGATGAGCTGTGAGGCTGTGCTCTCCGACGGA 395  
Qy 301 GGTAGCAGCTGGGCCACTCGAAACAGGACAGAGGCGCCCTCGTGTGCTCCGCTCCGCC 360  
Db 396 GGTAGCAGCTGGGCCACTCGAAACAGGACAGAGGCGCCCTCGTGTGCTCCGCTCCGCC 455  
Qy 361 ATCTCCAGGACGAGTGCAGATCCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 420

Db 456 ATCTCCAGGACGAGTGCAGCATCCAGAGCTGGAGCAGAGCTGTTGGAGAGGAGGCG 515  
Qy 421 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGAGAGAGAGCTTGCC 465  
Db 516 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGAGAGAGAGCTTGCC 560

## RESULT 6

US-10-091-548-17

; Sequence 17, Application US/10091548

; Publication No. US20030049703A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PA122C1

; CURRENT APPLICATION NUMBER: US/10/091,548

; CURRENT FILING DATE: 2002-03-07

; NUMBER OF SEQ ID NOS: 137

; Prior Application removed - See File Wrapper or Palm

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 17

; LENGTH: 560

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-091-548-17

Query Match 57.6%; Score 463.4; DB 9; Length 560;  
Best Local Similarity 99.8%; Pred. No. 3.6e-104;  
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GCCATCTGCACTCTCTCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGAC 60  
Db 96 GCCATCTGCACTCTCTCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGAC 155  
Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 120  
Db 156 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 215  
Qy 121 CGGAATCCATGTCCAGCTCCACACACAGCAGCAGCAGCTACCACTGAGACCCG 180  
Db 216 CGGAATCCATGTCCAGCTCCACACACAGCAGCAGCAGCTACCACTGAGACCCG 275  
Qy 181 CTGCTACACCCGTGGAGCCACAGCGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 240  
Db 276 CTGCTACACCCGTGGAGCCACAGCGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGC 335  
Qy 241 GGCATCGTCTCCAGGACAGCAATGATGAGCTGTGAGGCTGTGCTCTCCGACGGA 300  
Db 336 GGCATCGTCTCCAGGACAGCAATGATGAGCTGTGAGGCTGTGCTCTCCGACGGA 395  
Qy 301 GGTAGCAGCTGGGCCACTCGAAACAGGACAGAGGCGCCCTCGTGTGCTCCGCTCCGCC 360  
Db 396 GGTAGCAGCTGGGCCACTCGAAACAGGACAGAGGCGCCCTCGTGTGCTCCGCTCCGCC 455  
Qy 361 ATCTCCAGGACGAGTGCAGATCCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 420  
Db 456 ATCTCCAGGACGAGTGCAGATCCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGCG 515  
Qy 421 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGAGAGAGAGCTTGCC 465  
Db 516 GCCCTCCAGAGCTGCAGCGCAGCTTTGAGAGAGAGAGCTTGCC 560

## RESULT 7

US-10-074-095-269

; Sequence 269, Application US/10074095

; Publication No. US2003007704A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC008C1

; CURRENT APPLICATION NUMBER: US/10/074,095

; CURRENT FILING DATE: 2002-02-14

;; PRIOR APPLICATION NUMBER: 09/764,860  
;; PRIOR FILING DATE: 2001-01-17  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: 60/214,886  
;; PRIOR FILING DATE: 2000-06-28  
;; PRIOR APPLICATION NUMBER: 60/217,487  
;; PRIOR FILING DATE: 2000-07-11  
;; PRIOR APPLICATION NUMBER: 60/225,758  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/220,963  
;; PRIOR FILING DATE: 2000-07-26  
;; PRIOR APPLICATION NUMBER: 60/217,496  
;; PRIOR FILING DATE: 2000-07-11  
;; PRIOR APPLICATION NUMBER: 60/225,447  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/218,290  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: 60/225,757  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/226,868  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/216,647  
;; PRIOR FILING DATE: 2000-07-07  
;; PRIOR APPLICATION NUMBER: 60/225,267  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/216,880  
;; PRIOR FILING DATE: 2000-07-07  
;; PRIOR APPLICATION NUMBER: 60/225,270  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/251,869  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/235,834  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: 60/234,274  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: 60/234,223  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: 60/228,924  
;; PRIOR FILING DATE: 2000-08-30  
;; PRIOR APPLICATION NUMBER: 60/224,518  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/236,369  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/224,519  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/220,964  
;; PRIOR FILING DATE: 2000-07-26  
;; PRIOR APPLICATION NUMBER: 60/241,809  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/249,299  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/236,327  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/241,785  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/244,617  
;; PRIOR FILING DATE: 2000-11-01  
;; PRIOR APPLICATION NUMBER: 60/225,268  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/236,368  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/251,856  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/251,868  
;; PRIOR FILING DATE: 2000-12-08  
;; PRIOR APPLICATION NUMBER: 60/229,344  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/234,997  
;; PRIOR FILING DATE: 2000-09-25  
;; PRIOR APPLICATION NUMBER: 60/229,343  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,345  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,287  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: 60/229,513  
;; PRIOR FILING DATE: 2000-09-05  
;; PRIOR APPLICATION NUMBER: 60/231,413  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/229,509  
;; PRIOR FILING DATE: 2000-09-05  
;; PRIOR APPLICATION NUMBER: 60/236,367  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/237,039  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,038  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/236,370  
;; PRIOR FILING DATE: 2000-09-29  
;; PRIOR APPLICATION NUMBER: 60/236,802  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,037  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/237,040  
;; PRIOR FILING DATE: 2000-10-02  
;; PRIOR APPLICATION NUMBER: 60/240,960  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/239,935  
;; PRIOR FILING DATE: 2000-10-13  
;; PRIOR APPLICATION NUMBER: 60/239,937  
;; PRIOR FILING DATE: 2000-10-13  
;; PRIOR APPLICATION NUMBER: 60/241,787  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,474  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/246,532  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/249,216  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,210  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/226,681  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,759  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/225,213  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/227,182  
;; PRIOR FILING DATE: 2000-08-22  
;; PRIOR APPLICATION NUMBER: 60/225,214  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/235,836  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: 60/230,438  
;; PRIOR FILING DATE: 2000-09-06  
;; PRIOR APPLICATION NUMBER: 60/215,135  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: 60/225,266  
;; PRIOR FILING DATE: 2000-08-14  
;; PRIOR APPLICATION NUMBER: 60/249,218  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,208  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,213  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,212  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,207  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,245  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,244  
;; PRIOR FILING DATE: 2000-11-17



;; PRIOR APPLICATION NUMBER: 60/249,217  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,211  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,215  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,264  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,214  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/249,297  
;; PRIOR FILING DATE: 2000-11-17  
;; PRIOR APPLICATION NUMBER: 60/232,400  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/231,242  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,081  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/232,080  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,414  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/231,244  
;; PRIOR FILING DATE: 2000-09-08  
;; PRIOR APPLICATION NUMBER: 60/233,064  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/233,063  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,397  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,399  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/232,401  
;; PRIOR FILING DATE: 2000-09-14  
;; PRIOR APPLICATION NUMBER: 60/241,808  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,826  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,786  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/241,221  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/246,475  
;; PRIOR FILING DATE: 2000-11-08  
;; PRIOR APPLICATION NUMBER: 60/231,243  
;; PRIOR FILING DATE: 2000-09-08

Query Match 57.6%; Score 463.4; DB 9; Length 560;  
Best Local Similarity 99.8%; Pred. No. 3.6e-104;  
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60  
Db GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 155  
Qy 61 AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTCTGGGGCGCTGTCTGAGCTCCGGC 120  
Db AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTCTGGGGCGCTGTCTGAGCTCCGGC 215  
Qy 121 CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGCAGCTACAGCTGAGACCCG 180  
Db CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGCAGCTACAGCTGAGACCCG 275  
Qy 181 CTGGTTCACACCCGTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 240  
Db CTGGTTCACACCCGTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 335  
Qy 241 GGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTCTCTCCGACCGA 300  
Db GGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTCTCTCCGACCGA 395  
Qy 361 ATCTCCACGAGCTGCGGACATCCAGAGCTGAGAAAGAGCTGTTGGAGGAGGAGGC 420  
Db ATCTCCACGAGCTGCGGACATCCAGAGCTGAGAAAGAGCTGTTGGAGGAGGAGGC 275  
Qy 421 GGCCTCCAGAGCTGCGGACGCTTTGAGGAGGAGGAGCTTGGC 465  
Db GGCCTCCAGAGCTGCGGACGCTTTGAGGAGGAGGAGCTTGGC 335  
Qy 241 GGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTCTCTCCGACCGA 300  
Db GGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTCTCTCCGACCGA 395  
Qy 301 GGTAGCAAGCTGGGCCACTCGAAACAGGACAGCAAGGGCCCTCGTGTCTCCGCTCCCC 360

Db 396 GGTAGCAAGCTGGGCCACTCGAAACAGGACAGCAAGGGCCCTCGTGTCTCCGCTCCCC 455  
Qy 361 ATCTCCACGAGCTGCGGACATCCAGAGCTGAGAAAGAGCTGTTGGAGGAGGAGGC 420  
Db 456 ATCTCCACGAGCTGCGGACATCCAGAGCTGAGAAAGAGCTGTTGGAGGAGGAGGC 515  
Qy 421 GGCCTCCAGAGCTGCGGACGCTTTGAGGAGGAGGAGCTTGGC 465  
Db 516 GGCCTCCAGAGCTGCGGACGCTTTGAGGAGGAGGAGCTTGGC 560

## RESULT 8

US-09-764-860-269  
; Sequence 269, Application US/09764860  
; Patent No. US20020094953A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008  
; CURRENT APPLICATION NUMBER: US/09/764,860  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1198  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 269  
; LENGTH: 560  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-860-269

Query Match 57.6%; Score 463.4; DB 10; Length 560;  
Best Local Similarity 99.8%; Pred. No. 3.6e-104;  
Matches 464; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 60  
Db GCCATCTGCACTCTCTCCCGAGAGTGCAGCCACAGCTGCACCCCGCCCTCCAGAC 155

Qy 61 AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTCTGGGGCGCTGTCTGAGCTCCGGC 120  
Db AAGCCCAAGGAGCAGAGCTGAAGCCTGGCGCTGTCTGGGGCGCTGTCTGAGCTCCGGC 215  
Qy 121 CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGCAGCTACAGCTGAGACCCG 180  
Db CGGAATCCATGTCTCAGCTGCCACACAGCAGCAGCAGCTACAGCTGAGACCCG 275  
Qy 181 CTGGTTCACACCCGTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 240  
Db CTGGTTCACACCCGTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 335  
Qy 241 GGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTCTCTCCGACCGA 300  
Db GGCATCGTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTGTCTCTCCGACCGA 395  
Qy 301 GGTAGCAAGCTGGGCCACTCGAAACAGGACAGCAAGGGCCCTCGTGTCTCCGCTCCCC 360  
Db GGTAGCAAGCTGGGCCACTCGAAACAGGACAGCAAGGGCCCTCGTGTCTCCGCTCCCC 455  
Qy 361 ATCTCCACGAGCTGCGGACATCCAGAGCTGAGAAAGAGCTGTTGGAGGAGGAGGC 420  
Db ATCTCCACGAGCTGCGGACATCCAGAGCTGAGAAAGAGCTGTTGGAGGAGGAGGC 515  
Qy 421 GGCCTCCAGAGCTGCGGACGCTTTGAGGAGGAGGAGCTTGGC 465  
Db GGCCTCCAGAGCTGCGGACGCTTTGAGGAGGAGGAGCTTGGC 560

## RESULT 9

US-10-171-581-63  
; Sequence 63, Application US/10171581  
; Publication No. US20030104426A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue

; PRIOR FILING DATE: 2000-11-02  
 ; PRIOR APPLICATION NUMBER: US 60/290,029  
 ; PRIOR FILING DATE: 2001-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/290,645  
 ; PRIOR FILING DATE: 2001-05-15  
 ; PRIOR APPLICATION NUMBER: US 60/292,336  
 ; PRIOR FILING DATE: 2001-05-22  
 ; PRIOR APPLICATION NUMBER: US 60/295,798  
 ; PRIOR FILING DATE: 2001-06-06  
 ; PRIOR APPLICATION NUMBER: US 60/297,457  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,884  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,459  
 ; PRIOR FILING DATE: 2001-07-09  
 ; NUMBER OF SEQ ID NOS: 1740  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 194  
 ; LENGTH: 561  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA875041  
 ; US-09-917-800A-194  
  
 Query Match 14.0%; Score 112.4; DB 10; Length 561;  
 Best Local Similarity 77.0%; Pred. No. 1.8e-18;  
 Matches 137; Conservative 0; Mismatches 41; Indels 0;  
  
 QY 367 ACGGACGAGTGCAGCATCCAGGAGCTGGAACAGAGAAGCTGTTGGAGAGAGGAGGGCGGCTG  
 Db 561 ACGGATGAGTGCACCATTTCAGGAGCTGGAGCAGAAGCTGCTGCAGCGGAGACGTGCTG  
 QY 427 CAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCCTGGCCCTACG  
 Db 501 CAGAAGCTACAGCGCAGTTCGATGAGAAGGAGTTTGCCTCTGCGCCAGACCTTCTCG  
 QY 487 CGGCGCGCGCGCTGTCAGGCGCAGAGCTGAGAGGCGCCGAGGCCCAAGGCGGCAACG  
 Db 441 CGGCCACGCGGACTAGAGACGAACTGGAGTGCCTTGGAACTTAAGAGCAAGCTGCTG  
  
 RESULT 11  
 US-09-925-300-60  
 ; Sequence 60, Application US/09925300  
 ; Patent No. US20020151681A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Craig Rosen,  
 ; APPLICANT: Steve Ruben  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
 ; FILE REFERENCE: PA101  
 ; CURRENT APPLICATION NUMBER: US/09/925,300  
 ; CURRENT FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: PCT/US00/05988  
 ; PRIOR FILING DATE: 2000-03-08  
 ; PRIOR APPLICATION NUMBER: 60/124,270  
 ; PRIOR FILING DATE: 1999-03-12  
 ; NUMBER OF SEQ ID NOS: 1890  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 60  
 ; LENGTH: 1617  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1590)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1592)  
 ; OTHER INFORMATION: n equals a,t,g, or c  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1595)  
 ; OTHER INFORMATION: n equals a,t,g, or c

```

, PRIOR APPLICATION NUMBER: 60/124,270
, PRIOR FILING DATE: 1999-03-12
, NUMBER OF SEQ ID NOS: 1890
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 60
, LENGTH: 1617
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
, NAME/KEY: misc feature
, LOCATION: (1590)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (1592)
, OTHER INFORMATION: n equals a,t,g, or c
, NAME/KEY: misc feature
, LOCATION: (1595)
, OTHER INFORMATION: n equals a,t,g, or c

```

```

Qy 686 GGTCTTACGAGAGGGAGAACACCGCTTCGGCCCCCGCGCTGGAGGAG 732
Db 1270 AGCAGCAGGAGAGGACGAGCTGGAGCACCTGGAGCAGAGGTGGGGCAG 1316

RESULT 13
US-09-563-728A-31
; Sequence 31, Application US/09563728A
; Publication No. US20030078216A1
; GENERAL INFORMATION:
; APPLICANT: MacLeod, Alan R
; APPLICANT: Li, Zoumei
; APPLICANT: Besterman, Jeffrey M
; TITLE OF INVENTION: Inhibition of Histone Deacetylase
; FILE REFERENCE: 106101.129
; CURRENT APPLICATION NUMBER: US/09/563,728A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/132,287
; PRIOR FILING DATE: 1999-05-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 8459
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-563-728A-31

Query Match 7.2%; Score 57.8; DB 9; Length 8459;
Best Local Similarity 48.3%; Pred. No. 4.3e-05;
Matches 161; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 401 AGCTGTTGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGC 460
Db 980 AGCGGCCCTGCGGGAGCAGCAGCTGCAGCAGGAGCTCTCTGGCGCTCAAGCAGAAGCAGC 1039
Qy 461 TTGCTTCCAGCTGGCGCTACGAGGAGCGCGCGCGCTGTCAGGAGCAGAGCTGAGGGGCC 520
Db 1040 AGATCCAGAGGCAGATCTCTCATCGCTAGTTCCAGAGGAGCAGCAGCAGAGCTCTCCCGC 1099
Qy 521 CGAGAGCCAAAGCGCGCAACAGCTCAAGCAGCGCTCGCAGAGAGAGCCAGCGCGCGCAGC 580
Db 1100 AGCAGAGCGCGAGCTCCACGAGCACAATCAAGCAATAACAGGAGATGCTGGCCATGAAGC 1159
Qy 581 AGGTCTTCACCTGTCAGGTACTGTCAGCTTTCAGCAGGAGAGAGCGGCGAGCTCCGCGCAGGAGC 640
Db 1160 ACCAGCAGGAGCTGCTGGAAACACACAGCGGAGCTGGAGAGGCACCCCGCAGGAGCAGGAGC 1219
Qy 641 TCGAGAGCCTCATGAAGGAGCAGGAGCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGG 700
Db 1220 TGGAGAAGCAGCACCGGGAGCAGAAAGCTGCAGCAGCTCAAGAAACAAGGAGAGGGCAAG 1279
Qy 701 AGAAGACCAGCTTCGGCCCGCGCTCGAGAGA 733
Db 1280 AGAGTCCGTGGCCAGCAGAGAAGTGAAGATGA 1312

RESULT 14
US-09-817-913-8
; Sequence 8, Application US/09817913
; Patent No. US20020061860A1
; GENERAL INFORMATION:
; APPLICANT: Li, Zuomei
; APPLICANT: Bonfils, Claire
; APPLICANT: Besterman, Jeffrey
; TITLE OF INVENTION: Inhibition of Specific Histone Deacetylase Isoforms
; FILE REFERENCE: 106101.145
; CURRENT APPLICATION NUMBER: US/09/817,913
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/192,157
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8

```

Query Match 7.2%; Score 57.8; DB 10; Length 8459;  
Best Local Similarity 48.3%; Pred. No. 4.3e-05;  
Matches 161; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 401 AGCTGTTGGAGAGGAGGCGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGAGGAGC 460  
Db 980 AGCCGGCCCTGCGGAGCAGCAGCTGCAGGAGGAGCTCTGCGCTCAAGCAGAGCAGC 1039

Qy 461 TTGCTTCAGCCTGCGCTACAGGAGGCGCGCGCTGAGGAGCAGCTGAGGAGC 520  
Db 1040 AGATCCAGAGGAGAGTCTCTATCGCGGAGTTCCAGAGGAGCAGCAGAGCTCTCCCGGC 1099

Qy 521 CGGAGCCCAAGGCGGCAACAGCTCAAGCAGGCTCGCAGAGAGCCAGCGCGCAGC 580  
Db 1100 AGCAGAGGCGAGCTCCAGGACATCAAGCANTACAGAGATGCTGGCCATGAAGC 1159

Qy 581 AGGTCCTGACCTGACGCTGCTGAGTTCAGCAGGAGAGCGGCGAGCTCCGGCAGGAGC 640  
Db 1160 ACCAGCAGAGCTGCTGGAACACCAAGCAGGAGCTGAGAGGCGCCCGCAGGAGCAGC 1219

Qy 641 TCGAGAGCTCATGAGGAGGAGCAGCCTGCTGGAGACCAAGCTCAGGTCCTTACGAGAGG 700  
Db 1220 TGGAGAGCAGCAGCGGAGCAGAGCTGCGAGCAGCTCAAGAACAGGAGAGGCAAG 1279

Qy 701 AGAAGACAGCTTTCGGCCCCCGCGCTGGAGGAGA 733  
Db 1280 AGAGTGGCGTGGCCAGCAGAGTGAAGATGA 1312

Search completed: June 15, 2003, 06:49:12  
Job time : 119.039 secs

Qy 581 AGGTCCTGACCTGACGCTGCTGAGTTCAGCAGGAGAGCGGCGAGCTCCGGCAGGAGC 640  
Db 1160 ACCAGCAGAGCTGCTGGAACACCAAGCAGGAGCTGAGAGGCGCCCGCAGGAGCAGC 1219

Qy 641 TCGAGAGCTCATGAGGAGGAGCAGCCTGCTGGAGACCAAGCTCAGGTCCTTACGAGAGG 700  
Db 1220 TGGAGAGCAGCAGCGGAGCAGAGCTGCGAGCAGCTCAAGAACAGGAGAGGCAAG 1279

Qy 701 AGAAGACAGCTTTCGGCCCCCGCGCTGGAGGAGA 733  
Db 1280 AGAGTGGCGTGGCCAGCAGAGTGAAGATGA 1312

RESULT 15  
US-09-817-538-8  
; Sequence 8, Application US/09817538  
; Patent No. US20020137162A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Zuomei  
; APPLICANT: Bonfils, Claire  
; APPLICANT: Besterman, Jeffrey  
; TITLE OF INVENTION: Antisense Oligonucleotide Inhibition of Specific Histone  
; FILE REFERENCE: 106101.144  
; CURRENT APPLICATION NUMBER: US/09/817,538  
; CURRENT FILING DATE: 2001-03-26  
; PRIOR FILING DATE: 2000-03-24  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 8459  
; TYPE: DNA  
; ORGANISM: Human  
US-09-817-538-8

Query Match 7.2%; Score 57.8; DB 10; Length 8459;  
Best Local Similarity 48.3%; Pred. No. 4.3e-05;  
Matches 161; Conservative 0; Mismatches 172; Indels 0; Gaps 0;

Qy 401 AGCTGTTGGAGAGGAGGCGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGAGGAGC 460  
Db 980 AGCCGGCCCTGCGGAGCAGCAGCTGCGAGGAGCTCTGCGCTCAAGCAGAGCAGC 1039

Qy 461 TTGCTTCAGCCTGCGCTACAGGAGGCGCGCGCTGAGGAGCAGCTGAGGAGC 520  
Db 1040 AGATCCAGAGGAGAGTCTCTATCGCGGAGTTCCAGAGGAGCAGCAGAGCTCTCCCGGC 1099

Qy 521 CGGAGCCCAAGGCGGCAACAGCTCAAGCAGGCTCGCAGAGAGCAGCGCGCGCAGC 580  
Db 1100 AGCAGAGGCGCAGCTCCAGCAGCAGCATCAAGCAATAACAGAGAGTCTGGCCATGAAGC 1159

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 1119.4 Seconds  
(without alignments)  
11632.279 Million cell updates/sec

Title: US-09-513-888c-1\_COPY\_1707\_2510

Perfect score: 804

Sequence: 1 gccatctgactctcccc.....tttggccagtagaccacct 804

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estmu:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vit:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_othr:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	507.2	63.1	740	14 BQ769435	BQ769435 UI-M-FIO-
2	468.6	58.3	547	12 BF590813	BF590813 7h42e07.x
3	463.4	57.6	545	10 AW007737	AW007737 wt68e06.x
4	448.8	55.8	668	12 BF058214	BF058214 7k28q05.x
5	445.4	55.4	447	9 A1042490	A1042490 ox62e04.x
6	429.8	53.5	817	9 A1984777	A1984777 wr85b07.x

C	7	365	45.4	397	10	AW028197	AW028197
C	8	336.6	41.9	344	10	AW016544	AW016544 UI-H-BIOP
C	9	296.8	36.9	743	10	BE410921	BE410921 601303579
C	10	285.2	35.5	673	10	BE384131	BE384131 601272956
C	11	254.8	31.7	407	10	AV663617	AV663617 AV663617
C	12	246	30.6	394	12	BF549120	BF549120 UI-R-AO-a
C	13	245	30.5	394	12	BF554804	BF554804 UI-R-E0-c
C	14	179	22.3	725	9	AJ454524	AJ454524 AJ454524
C	15	172.4	21.4	531	10	BE276168	BE276168 601144424
C	16	151.4	18.8	153	9	AA644373	AA644373 af64c10.8
C	17	144.4	18.0	596	10	AV663618	AV663618 AV663618
C	18	135.8	16.9	926	12	BF583930	BF583930 602098941
C	19	125.4	15.6	907	14	BQ652760	BQ652760 AGENCOURT
C	20	125.4	15.6	921	14	BQ652556	BQ652556 AGENCOURT
C	21	123.8	15.4	991	14	BQ720762	BQ720762 AGENCOURT
C	22	123.4	15.3	941	14	BQ955592	BQ955592 AGENCOURT
C	23	121.2	15.3	635	-12	BF031137	BF031137 601456991
C	24	115	14.3	924	12	BE897837	BE897837 601440570
C	25	113.8	14.2	940	14	BQ895397	BQ895397 AGENCOURT
C	26	112.4	14.0	561	9	AA875041	AA875041 UI-R-E0-c
C	27	112.2	14.0	1101	14	BQ051027	BQ051027 AGENCOURT
C	28	109.8	13.7	460	12	BE757942	BE757942 212520 MA
C	29	106.2	13.2	647	12	BG338379	BG338379 602436164
C	30	105.4	13.1	700	13	B1854892	B1854892 603381873
C	31	104.6	13.0	663	9	AL635635	AL635635 AL635635
C	32	104.6	13.0	1115	12	BF784037	BF784037 602107882
C	33	104.4	13.0	735	12	BE740090	BE740090 601595016
C	34	104.2	13.0	906	13	BG917774	BG917774 602818743
C	35	100.6	12.5	487	13	B1338021	B1338021 361664 MA
C	36	99	12.3	1093	14	BM811378	BM811378 AGENCOURT
C	37	98.6	12.3	661	12	BF312957	BF312957 601896225
C	38	94	11.7	602	14	BM717087	BM717087 UI-E-EJO-
C	39	94	11.7	603	14	BM674468	BM674468 UI-E-EJO-
C	40	92.4	11.5	723	10	BE312971	BE312971 601150210
C	41	90.4	11.2	806	14	BM947278	BM947278 UI-M-SHOP
C	42	89.6	11.1	932	14	BQ879505	BQ879505 AGENCOURT
C	43	87.8	10.9	846	12	BF970582	BF970582 602274088
C	44	87.8	10.9	923	12	BG750395	BG750395 602709229
C	45	87.8	10.9	1298	13	BM470160	BM470160 AGENCOURT

#### ALIGNMENTS

RESULT 1  
BQ769435  
LOCUS BQ769435 740 bp mRNA linear EST 26-JUL-2002  
DEFINITION UI-M-FIO-byq-a-22-0-UI.r1 NIH\_BMAP\_FIO Mus musculus cDNA clone  
IMAGE:6400797 5', mRNA sequence.  
ACCESSION BQ769435  
VERSION BQ769435.1 GI:21977909  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 740)  
AUTHORS NTH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Cloned through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
Seq primer: pYX-5.  
Location/Qualifiers

#### FEATURES





/tissue type="fibrotheoma"  
/lab host="DH10B (phage-resistant)"  
/note="organ: ovary; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGGACATTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 185 c 226 g 150 t 2 others  
ORIGIN  
Query Match 55.8%; Score 448.8; DB 12; Length 668;  
Best Local Similarity 99.3%; Pred. No. 8.9e-78;  
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GCATCTCTGCACTCTCCCGGAGAGTGCAGCCACACAGCTGCACCCGCGCCCTCCAGAC 60  
DB 453 GCCATCTCTGCACTCTCCCGGAGAGTGCAGCCACACAGCTGCACCCGCGCCCTCCAGAC 394  
QY 61 AAGCCCAAGGAGCAGAGCTGAAGCCTGCTGTGCTCTGGGGCGCTGTCTGAGCTCCGGC 120  
DB 393 AAGCCCAAGGAGCAGAGCTGAAGCCTGCTGTGCTCTGGGGCGCTGTCTGAGCTCCGGC 334  
QY 121 CGGAATCTCATGTCTCAGCTGCTCCACACACAGCAGCAGCAGCTACAGCTGAGCCG 180  
DB 333 CGGAATCTCATGTCTCAGCTGCTCCACACACAGCAGCAGCAGCTACAGCTGAGCCG 274  
QY 181 CTGCTCACACCCGCTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 240  
DB 273 CTGCTCACACCCGCTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATCACCCAG 214  
QY 241 GCATCTCTCCAGGACGACCAATGATGAGCTTGAAGGCTGTGCTTTCGAGCGGA 300  
DB 213 GCGATCTGCTCTCCAGGACGACCAATGATGAGCTTGAAGGCTGTGCTTCTCCGACGGA 154  
QY 301 GGTACGAGCTGGGCACTCCGACAGCAGCAGAGGCGCCCTCGTGTGCTCGCTCCCC 360  
DB 153 GGTACGAGCTGGGCACTCCGACAGCAGCAGAGGCGCCCTCGTGTGCTCGCTCCCC 94  
QY 361 ATCTCCAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 420  
DB 93 ATCTCCAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 34  
QY 421 GCCCTCCAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 453  
DB 33 GCCCTCCAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT

RESULT 5  
AI042490/c  
LOCUS  
DEFINITION  
ox62e04.x1 Soares\_NHMPu\_S1 Homo sapiens cDNA clone IMAGE:1660926  
3', mRNA sequence.

ACCESSION  
AI042490  
VERSION  
AI042490.1 GI:3281684  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 447)

REFERENCE  
AUTHORS  
TITLE  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
Insert Length: 656 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 435.  
FEATURES  
Location/Qualifiers  
source  
1..447  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone="IMAGE:1660926"  
/clone\_lib="Soares\_NHMPu\_S1"  
/tissue\_type="Pooled human melanocyte, fetal heart, and pregnant uterus"  
/lab host="DH10B"  
/note="Organ: mixed (see below); Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NHMP, pregnant uterus NHMPu, and fetal heart NHMPu) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. Clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 62 a 133 c 156 g 96 t  
ORIGIN  
Query Match 55.4%; Score 445.4; DB 9; Length 447;  
Best Local Similarity 99.8%; Pred. No. 3.8e-77;  
Matches 446; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 19 CCGAGAGTGCACGCCACACAGCTGCACCCGCGCCCTCCAGACAGCCCAAGGAGCAGGAG 78  
DB 447 CCGAGAGTGCACGCCACACAGCTGCACCCGCGCCCTCCAGACAGCCCAAGGAGCAGGAG 388  
QY 79 CTGAGGCTGGCTGTCTCTGGGGCGCTGTCTGAGCTCCGGCCGGAACCTCATGTCCAGC 138  
DB 387 CTGAAGCTGGCTGTCTCTGGGGCGCTGTCTGAGCTCCGGCCGGAACCTCATGTCCAGC 328  
QY 139 CTGCCCCACACAGCAGCAGCAGCTACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 198  
DB 327 CTGCCCCACACAGCAGCAGCAGCTACAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 268  
QY 199 CCCAACAGCGCTTTGGGGCTCCGCCCAACATCACCCAGGCGCATCTCTCCAGGAC 258  
DB 267 CCCAACAGCGCTTTGGGGCTCCGCCCAACATCACCCAGGCGCATCTCTCCAGGAC 208  
QY 259 AGCAATATGATGAGCTTGAAGCTCTCTCTTCTCCAGGAGGTAGCAAGCTGGGCGAC 318  
DB 207 AGCAATATGATGAGCTTGAAGCTCTCTCTTCTCCAGGAGGTAGCAAGCTGGGCGAC 148  
QY 319 TCGAACAGGAGCAGACAGGCGCCCTCGTGTGCTCGCTCCCGCATCTCCAGGAGGAGTGC 378  
DB 147 TCGAACAGGAGCAGACAGGCGCCCTCGTGTGCTCGCTCCCGCATCTCCAGGAGGAGTGC 88  
QY 379 AGCATCCAGGAGCTGGGAACAGAGCTCTTGGAGAGGAGGCGCGCCCTCCAGAGCTGCAG 438  
DB 87 AGCATCCAGGAGCTGGGAACAGAGCTCTTGGAGAGGAGGCGCGCCCTCCAGAGCTGCAG 28  
QY 439 CGCAGCTTTGAGGAGGAGGAGCTTGCC 465  
DB 27 CGCAGCTTTGAGGAGGAGGAGCTTGCC 1

RESULT 6  
AI984777/c  
LOCUS  
DEFINITION  
AI984777  
817 bp mRNA linear EST 08-MAR-2000  
similar to TR:O60299 O60299 KIAA0552 PROTEIN.; mRNA sequence.  
ACCESSION  
AI984777  
VERSION  
AI984777.1 GI:5812054  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 817)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 1307 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 428.

## FEATURES

Location/Qualifiers  
1..817

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2494453"

/clone\_lib="NCI-CGAP\_Kid11"

/lab\_host="DH10B"

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI-CGAP Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1323176-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 133 a 235 c 266 g 180 t 3 others

Query Match 53.5%; Score 429.8; DB 9; Length 817;  
Best Local Similarity 97.6%; Pred. No. 4.7e-74;  
Matches 457; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy 1 GCCATCTGCACTCTCCCGGAGAGTCCAGCCACACAGCTGCACCCCGCCCTCCAGAC 60  
Db GCCATCTGCACTCTCCCGGAGAGTCCAGCCACACAGCTGCACCCCGCCCTCCAGAC 409  
Qy 61 AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGGCTGTCTAGACTCCGGC 120  
Db AAGCCCAAGGAGCAGGAGCTGAAGCTGGCTGTCTCTGGGGGCTGTCTAGACTCCGGC 349  
Qy 121 CGGAATCTCATGTCTCAGCTGCCACACACAG-CACCAAGCAGCAGCTACCAAGCT--GGAC 177  
Db CGGAATCTCATGTCTCAGCTGCCACACACAGCCCCCAGCAAGATTACCAAGTTGGGAC 289  
Qy 178 CCGTGTGTACACCCGTTGGGACCCCAAGCCGTTTGGGGGCTCGGCCCAACAATCACC 237  
Db CCGTGTGTACACCCGTTGGGACCCCAAGCCGTTTGGGGGCTCGGCCCAACAATCACC 229  
Qy 238 CAGGCGATCGTCTCCAGCAGCAGCAATGATGAGCTTGAAGGCTCTGTCTTCTCCGAC 297  
Db CAGGCGATCGTCTCTCCAGCAGCAGCAATGATGAGCTTGAAGGCTCTGTCTTCTCCGAC 169  
Qy 298 GGAGGTAGCAAGCTGGGGCACTCGAAACAAGGCAGACAAGGGCCCTCGTGTGTCCGCTCC 357  
Db GGAGGTAGCAAGCTGGGGCACTCGAAACAAGGCAGACAAGGGCCCTCGTGTGTCCGCTCC 109  
Qy 358 CCCATCTCCAAGGAGTGCAGCATCCAGAGCTGGAAACAGAGCTGTTGGAGGGAG 417  
Db CCCATCTCCAAGGAGTGCAGCATCCAGAGCTGGAAACAGAGCTGTTGGAGGGAG 49  
Qy 418 GGCGCCCTCCAGAGCTCAGCGGAGCTTTGAGGAGAGGAGCTTGCC 465

Db 48 GCGCGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCC 1

## RESULT 7

AW028197/c

LOCUS

DEFINITION

AW028197

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW028197

WY84A01.X1 Soares\_thymus\_NHFT Homo sapiens cDNA clone

IMAGE:2536200 3' Similar to TR:O60299 O60299 KIAA0552 PROTEIN. ;

mRNA sequence.

AW028197

AW028197.1 GI:5885953

EST.

human.

Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 397)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Insert Length: 1067 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 139.

Location/Qualifiers

1..397

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2536200"

/clone\_lib="Soares\_thymus\_NHFT"

/dev\_stage="fetal"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: thymus, pooled; Vector: pT7T3D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5',

TGTACCAATCTGAGTGGAGCGCGCCGACGTTTCTTTTCTTTTCTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 67 a 120 c 124 g 86 t

ORIGIN

Query Match 45.4%; Score 365; DB 10; Length 397;

Best Local Similarity 95.0%; Pred. No. 1.7e-61;

Matches 377; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 97 TCTGGGCGCTGTCTCAGACTCCGGCGGAATCTCCATGTCCAGCTGCCACACAGCACC 156

Db TCTGGGCGCTGTCTCAGACTCCGGCGGAATCTCCATGTCCAGCTGCCACACAGCACC 338

Qy 157 AGCAGCAGCTACCAAGCTGGAGCCGCTGTCTCACACCCCTGGGACCCCAAGCCGTTTGGG 216

Db GGCAGCAGCTACCAAGCTGGAGCCGCTGTCTCACACCCCTGGGACCCCAAGCCGTTTGGG 278

Qy 217 GGCTCCGCCCAACAATCATCACCAGGGCATCGTCTCCAGGACAGCAACATGATGAGCCTG 276

Db GGCTCCGCCCAACAATCATCACCAGGGCATCGTCTCCAGGACAGCAACATGATGAGCCTG 218

Qy 277 AAGCTCTGTCTTCTTCTCGAGCGAGGTAGCAGCTGGGCCCACTCGAACAGGACAGCAAG 336

Db AAGCTCTGTCTTCTTCTCGAGCGAGGTAGCAGCTGGGCCCACTCGAACAGGACAGCAAG 158

Qy 337 GGCCCTCTGTGTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 396

Db GGCCCTCTGTGTGTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 98

QY 397 CAGAGCTGTTGGAGAGGAGGGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGAAG 456  
 |||||  
 Db 97 CAGAGCTGTTGGAGAGGAGGGCCCTCCAGAGCTGCAGCGAGCTTTGAGGAGAAG 38  
 |||||  
 QY 457 GAGCTTGCTCCAGCTCGCTTACGAGGAGCGCGC 493  
 |||||  
 Db 37 GAGCTTGCTCCAGCTCGCTTACGAGGAGCGCGC 1

RESULT 8  
 AW016544/c  
 LOCUS  
 DEFINITION  
 UI-H-B10p-abg-g-04-0-UI.s1 NCI\_CGAP\_Sub2 Homo sapiens cDNA clone  
 IMAGE:2711982 3', mRNA sequence.  
 ACCESSION  
 AW016544  
 VERSION  
 AW016544.1 GI:5865301  
 KEYWORDS  
 EST.  
 SOURCE  
 Human.

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 344)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov

Oligo-dT track not found, Not 1 site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: NCI-CGAP clone distribution  
 information can be found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Seq primer: M13 Forward  
 POLYA=No.

#### FEATURES

Location/Qualifiers  
 1. 344  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2711982"  
 /clone\_lib="NCI CGAP Sub2"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pVT30-Pac (Pharmacia) with a modified  
 polylinker. Site 1: Not 1; Site 2: Eco RI; The  
 NCI\_CGAP\_Sub2 library is a subtracted library derived from  
 B1. B1 constitutes a mixture of 21 normalized or  
 subtracted NCI\_CGAP libraries: NCI\_CGAP\_Co4, NCI\_CGAP\_Pr22,  
 NCI\_CGAP\_Pr28, NCI\_CGAP\_Co10, NCI\_CGAP\_Co16,  
 NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12, NCI\_CGAP\_Co16,  
 NCI\_CGAP\_Kid11, NCI\_CGAP\_Lym2, NCI\_CGAP\_Br2, NCI\_CGAP\_Co8,  
 NCI\_CGAP\_Co11, NCI\_CGAP\_Le12, NCI\_CGAP\_Brn23, NCI\_CGAP\_Lu5,  
 NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6,  
 NCI\_CGAP\_Brn25. These 21 libraries were pooled and a  
 single-stranded DNA preparation of the resulting mixture  
 was used as a tracer in a subtractive hybridization with a  
 driver whose composition is detailed below: NCI\_CGAP\_Kid3  
 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE\_  
 Clonids 132376-132391), 1456008-1456775, 1500552-1502855  
 ) NCI\_CGAP\_Kid5 pool 1 LLAM 3338-3342, 3722-3725,  
 3776-3778 (IMAGE Clonids 1323912-1325831,  
 1471368-1472903, 1492104-1493255) NCI\_CGAP\_Lu5 pool 1 LLAM  
 3575-3582, 3851-3854 (IMAGE Clonids 1414920-1417991,  
 1520904-1522439) NCI\_CGAP\_GC4 pool 1 LLAM 3164-3167,  
 3716-3720, 3733-3735 (IMAGE Clonids 1257096-1258631,  
 1469064-1470983, 1475592-1476743) NCI\_CGAP\_Pr22 pool 1  
 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonids  
 985608-986759, 1101192-1101959, 1217928-1220615)  
 NCI\_CGAP\_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE  
 Clonids 1057416-1061255, 1144584-1145351) The resulting  
 subtracted library contained 4 million recombinants.  
 Subtraction was performed as previously described (Bonaldi  
 & Soares (1996): Normalization and Subtraction:  
 Two Approaches To Facilitate Gene Discovery. Genome

Research 6, 791-806.  
 TAG\_L1B=NCI CGAP GC4  
 TAG\_TISSUE=germ cell  
 TAG\_SEQ=AAATC"

BASE COUNT 51 a 106 c 112 g 74 t 1 others  
 ORIGIN

Query Match 41.9%; Score 336.6; DB 10; Length 344;  
 Best Local Similarity 99.5%; Pred. No. 5.8e-56;  
 Matches 339; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 149 ACAGCACCAGCAGCAGCTACAGCTGGACCGCTGGTTCACACCCGTCGGGACCCCAAGCC 208  
 |||||  
 Db 344 ACAGCACCAGCAGCAGTTACAGCTGGACCGCTGGTTCACACCCGTCGGGACCCCAAGCC 285  
 |||||  
 QY 209 GTTTTGGGGGCTCGCCCAACAATCAACCAGGCGATCGTCTCCAGGACGACGAACATGA 268  
 |||||  
 Db 284 GTTTTGGGGGCTCGCCCAACAATCAACCAGGCGATCGTCTCCAGGACGACGAACATGA 225  
 |||||  
 QY 269 TGAGCCTGAAGGCTCTGTCTTCCGACGAGGTAGCAAGCTGGGCGCACTCCGAACAAGG 328  
 |||||  
 Db 224 TGAGCCTGAAGGCTCTGTCTTCCGACGAGGTAGCAAGCTGGGCGCACTCCGAACAAGG 165  
 |||||  
 QY 329 CAGACAAGGGCCCTCGTGTGTCGGCTCCGCCCATCTCCAGGACGAGTGCAGCATCCAGG 388  
 |||||  
 Db 164 CAGACAAGGGCCCTCGTGTGTCGGCTCCGCCCATCTCCAGGACGAGTGCAGCATCCAGG 105  
 |||||  
 QY 389 AGCTGGAACAGAGCTTTGGAGAGGAGGGCGGCCCTCCAGAGCTCGAGCGGAGCTTTG 448  
 |||||  
 Db 104 AGCTGGAACAGAGCTTTGGAGAGGAGGGCGGCCCTCCAGAGCTCGAGCGGAGCTTTG 45  
 |||||  
 QY 449 AGGAGAAGGAGCTTGGCTTCAGCCTCGCTTACGAGGAGCGCGC 492  
 |||||  
 Db 44 AGGAGAAGGAGCTTGGCTTCAGCCTCGCTTACGAGGAGCGCGC 1

#### RESULT 9

BE410921/c  
 LOCUS  
 DEFINITION  
 601303579F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3638203 5',  
 mRNA sequence.  
 ACCESSION  
 BE410921  
 VERSION  
 BE410921.1 GI:9347371  
 KEYWORDS  
 EST.  
 SOURCE  
 human.

ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 743)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LLCM338 row: m column: 20  
 High quality sequence start: 49  
 High quality sequence stop: 688.  
 Location/Qualifiers

#### FEATURES

1. 743  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3638203"  
 /clone\_lib="NIH\_MGC\_21"  
 /tissue\_type="choriocarcinoma"  
 /lab\_hosts="DH10B (phage-resistant)"  
 /note="Organ: Placenta; Vector: pORF7; Site 1: XhoI;  
 Site 2: EcoRI; cDNA made by oligo-dT priming."



```

/clone="E1BR037F07"
/clone_lib="Bos taurus brain fetus"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/notes="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A
was deleted from a NotI site"
BASE COUNT      58 a   119 c   141 g   89 t
ORIGIN

Query Match      31.7%; Score 254.8; DB 10; Length 407;
Best Local Similarity 84.8%; Pred. No. 5.4e-40;
Matches 324; Conservative 0; Mismatches 47; Indels 11; Gaps 3;

QY 64 CCCAAGGAGCAGGAGCTGAAGCTGGCTGTGCTTGGGGCG--CTGTCAAGCTCCGGCC 121
Db 373 CCAAAGGAGCAGGAGCTGAAGCCAGCCTGTGCTCCGGGGCGCACTGTCCGAACCTCCGGCC 314
QY 122 GGAATCTCATGTCCAGCTGCCACACACAGCAGCAGCAGCTACCAGCTGGACCGC 181
Db 313 GGAATCTCATGTCCAGCTGCCACACAGCAGCAGCAGCTACCAGCTGGACCGC 254
QY 182 TGGTCACACCCGTGGAGCCACCAAGCCGTTTGGGGGCTCCGCCACCAACATCACCCAGG 241
Db 253 TGGTCACCTCTGTGGGGCCGCGCCGCTTGGGGGCTCAGCCACCAACATCACACAGG 194
QY 242 GCATCGTCTCCAGACAGCAACATGATGAGCTGAAGCTGTGCTTCTCCGACGAG 301
Db 193 GCATTTGCTCTCAAGACAGCAACATGATGAGCTGAAGCGCTGTCTTCTCTGACGGG 134
QY 302 GTAGCAAGCTGGCCACTCGACAGCAGCAGCAGGCGCCCTCGTGTCTCGCTCCCGCA 361
Db 133 GCAGCAAGCTGGCCCGCCAGCAAGCAA---AGACAAGGGC-----TCTGTGGCTCCCGCA 83
QY 362 TCTTCAACGAGCAGTGCGACATCCAGGAGCTGGAACAGAACTGTGTGAGAGGAGGGCG 421
Db 82 TCTTCAACGAGCAGTGCGACATCCAGGAGCTGGAACAGAACTGTGTGAGAGGAGGGCG 23
QY 422 CCTTCCAGAGCTGAGGCGAG 443
Db 22 AGCTGCAGAGGCTGCACCGCAG 1

RESULT 12
BF549120
LOCUS
DEFINITION
UI-R-A0-af-c-04-0-UI.r1 UI-R-A0 Rattus norvegicus cdna clone
UI-R-A0-af-c-04-0-UI 5', mRNA sequence.
ACCESSION
BF549120
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 394)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
9704477
MEDLINE
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cdna Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1791781

Seq primer: M13 Forward.
Location/Qualifiers
1. .394
/organism="Rattus norvegicus"
/db_xref="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-A0-af-c-04-0-UI"
/clone_lib="UI-R-A0"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; This library
consists of a mixture of individually tagged normalized
libraries constructed from rat placenta, adult lung, brain
liver, kidney, heart, spleen, ovary, and muscle. The tag
is a string of 3-5 nucleotides present between the Not I
site and the oligo-dT track which allows identification of
the library of origin of a clone within the mixture."
BASE COUNT      89 a   125 c   103 g   77 t
ORIGIN

Query Match      30.6%; Score 246; DB 12; Length 394;
Best Local Similarity 85.8%; Pred. No. 2.8e-38;
Matches 273; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 GCCATCTGTCACTCTCCCGGAGAGTGCCAGCAGCAGCTGCACCCGCCCTCCAGAC 60
Db 76 GCCATCTGTCACTCTCCCGGAGAGTGCCAGCAGCAGCTGCACCCGCCCTCCAGAC 135
QY 61 AAGCCCAAGCAGGAGGAGCTGAGCCTGGCTGTGCTCTGGGGCGTGTGCAGACTCCGGC 120
Db 136 AAGCCCAAGCAGGAGGAGCTGAGCCTGGCTGTGCTCTGGGGCGTGTGCAGACTCCGGC 195
QY 121 CGGAATCTCATGTCTCAAGCTGCCACACAGCAGCAGCAGCTACAGCTGGACCCG 180
Db 196 CGGAATCTCATGTCTCAAGCTGCCACACAGCAGCAGCAGCTACAGCTGGATCCT 255
QY 181 CTGCTCACACCCGTGGAGCCCAAGCGCTTTGGGGGCTCCGCCACACATCACCCAG 240
Db 256 CTGCTCACACCCGTGGAGCCCAAGCGCTTACCTAGCCGTTTGGGGGTTTCAAGCTCTGCTGATGGG 315
QY 241 GGATCTCTCTCCAGGACAGCAACATGATGAGCTGAAGGCTCTGTCTTCTCCGAGGA 300
Db 316 GGATCTCTCTCCAGGACAGTAATATGATGAGCTGAAGGCTCTGTCTTCTGATGGG 375
QY 301 GGTAGCAAGCTGGGCCAC 318
Db 376 GGCAGCAAGCTGGGCTCAC 393

RESULT 13
BF554804
LOCUS
DEFINITION
UI-R-E0-cb-b-03-0-UI.r1 UI-R-E0 Rattus norvegicus cdna clone
UI-R-E0-cb-b-03-0-UI 5', mRNA sequence.
ACCESSION
BF554804
VERSION
KEYWORDS
SOURCE
EST.
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 394)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
9704477
MEDLINE
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA

```

Tel: 319 335 8250  
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu  
CDNA Library Preparation: M.B. Soares Lab Clone Distribution:  
clones will be available through Research Genetics (www.resgen.com)  
This clone is also available through the I.M.A.G.E. Consortium at  
LLNL (info@image.llnl.gov). IMAGE ID= 1770031  
Seq primer: M13 Forward.

FEATURES  
source  
Location/Qualifiers  
1..394

/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-E0-cb-b-03-0-UI"  
/clone\_lib="UI-R-E0"  
/dev\_stage="embryonic"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: NotI; Site\_2: EcoRI; This library  
consists of a mixture of individually tagged normalized  
libraries constructed from 8, 12 and 18-day embryo. The  
tag is a string of 3-5 nucleotides present between the  
Not I site and the oligo-dT track which allows  
identification of the library of origin of a clone within  
the mixture."

BASE COUNT 89 a 125 c 102 g 77 t 1-others

Query Match 30.5%; Score 245; DB 12; Length 394;

Best Local Similarity 85.5%; Pred. No. 4.3e-38;  
Matches 272; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 GCCATCTCCATCTCTCCCGGAGAGTGCACGACACACAGCTGACCCCGCCCTCCAGAC 60  
DB 76 GCCATCTCCATCTCTCCCGGAGAGTGCACGACACACAGCTGACCCCGCCCTCCAGAT 135  
QY 61 AAGCCCAAGGACGAGCTGAGCTGGCTGTGCTCTGGGGCGCTGTGACACTCCGGC 120  
DB 136 AAGCCCAAGGACGAGCTGAGCTGGCTGTGCTCTGGGGCGCTGTGACACTCCGGC 195  
QY 121 CGGAATCCATGCTCCAGCTGCCACACACAGCAGCAGCAGCTACCGCTGACGTCGACCG 180  
DB 196 CGGAATCCATGCTTAGCTGCCACGATAGCAGCAGCAGCTACCGCTGACGTCGATCCT 255  
QY 181 CTGCTCACACCGTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATACACCCAG 240  
DB 256 CTGCTCACACCGTGGGACCCACAGCCGTTTGGGGCTCCGCCCAACATACACCA 315  
QY 241 GGCATCGTCTCCAGGACGACGACATGATGAGCTGAAGGCTGTCTCTTCGACGGA 300  
DB 316 GGCATCATCTTCAGGACAGTAATATGATGAGCTGAAGGCTGTCTCTTCGATGGG 375  
QY 301 GGTAGCAAGCTGGGCCAC 318  
DB 376 GGCACCAAGCTGGCTCAC 393

RESULT 14

AJ454524

LOCUS

DEFINITION AJ454524 riken1 Gallus gallus cdna clone 3a8r1, mRNA sequence.

ACCESSION AJ454524

VERSION AJ454524.1 GI:20264620

KEYWORDS EST.

SOURCE

ORGANISM Gallus gallus

chicken.

REFERENCE

AJ454524

1 (bases 1 to 725)

Buerstedde, J.M.

Gallus gallus bursal lymphocyte EST

Unpublished (2002)

COMMENT

Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.

FEATURES

source

Location/Qualifiers  
1..725

/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="3a8r1"  
/clone\_lib="riken1"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/note="CB inbred strain"

BASE COUNT 138 a 190 c 293 g 100 t 4 others

Query Match 22.3%; Score 179; DB 9; Length 725;

Best Local Similarity 66.8%; Pred. No. 3.7e-25;

Matches 255; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 383 TCCAGGAGCTGGAACAGAGCTGTTGGAGAGGAGGCGCCCTCCAGAGCTGCAGCGCA 442  
DB 68 TCGTGAGCTGGAACAGAGCTGCGGAGAGGAGGCGGAGCTGCAGAGCTGCAGGACA 127  
QY 443 GCTTTGAGGAGAAGAGCTTGCCTCCAGCTGCGCTACGAGAGCGCGCGCGCTGCA 502  
DB 128 GCCTGGATGAGATGAGGTGGCCATCTGCCAGGTGTACGAAGAAGCAGCGTCTGTG 187  
QY 503 GGGAGAGCTGGAGGCGCGGAGCCCAAGGGGGCAACAAGCTCAAGCAGGCTCGCAGA 562  
DB 188 AGCAGGAGCTGGAGGCTTGGCGCAGCGCTGTGGCGCCAGGTGCGACAGCGGCCAGC 247  
QY 563 AGAGCCAGCGCGCAGCAGCTGCTGCACTGCGAGTACTGAGTCTCAGCAGGAGAGC 622  
DB 248 AGGCGCAGCTGGGAGCAGGCTGCTGCACTGCGAGTCTGCGAGTCTGCGAGTCTGCGAGGAGAGA 307  
QY 623 GGCAGCTCCGCGCAGGAGCTCGAGAGCTTCATGAAGGAGCAGGAGCTGCTGGAGACCAAGC 682  
DB 308 AGCAGCTGCGAGGAGCTTGTCTCAGCTGCTGCGAGGCGGAGCTGCTGGAGCGCGCT 367  
QY 683 TCAGGTCTCAGAGGAGGAGAACAGCTTCGCGCCCGCGCTGGAGAGACCCAGTGGG 742  
DB 368 CGCATCTCTCCAGCGGAAACACAGGAGCTGGGGCCACGCTGGAGGAGACCAAGTGGG 427  
QY 743 AGGTGAGGCCACACAGGCTCA 764  
DB 428 AGGTGCGCAGAAGTCCGGTGA 449

RESULT 15

BE276168/c

LOCUS

DEFINITION 60114424P1 NIH\_MGC\_20 Homo sapiens cdna clone IMAGE:3050716 5',

mRNA sequence.

ACCESSION BE276168

VERSION BE276168.1 GI:9151131

KEYWORDS EST.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 531)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

AUTHORS

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LCM96 row: 0 column: 05  
High quality sequence stop: 531.  
Location/Qualifiers  
1. 531  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3050716"  
/clone\_lib="NIH MGC\_20"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pOT87; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 87 a 173 c 156 g 115 t  
ORIGIN

Query Match 21.4%; Score 172.4; DB 10; Length 531;  
Best Local Similarity 99.4%; Pred. No. 6.6e-24;  
Matches 173; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 376 TGCAGCATCCAGGAGCTGGACAGAAAGCTGTGGAGAGGAGGGCGCCCTCCAGAAAGCTG 435  
|||  
Db 531 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTGGAGAGGAGGGCGCCCTCCAGAAAGCTG 472  
QY 436 CAGCGCAGCTTTGAGGAGAAGAGCTTGCTCCAGCCTGGCTACGAGGAGCGCGCGG 495  
|||  
Db 471 CAGCGCAGCTTTGAGGAGAAGAGCTTGCTCCAGCCTGGCTACGAGGAGCGCGCGG 412  
QY 496 CGCTCAGGAGCAGCTGGAGGGCCCGGAGCCCAAAGCGGCAACAAGCTCAAG 549  
|||  
Db 411 CGCTCAGGAGCAGCTGGAGGGCCCGGAGCCCAAAGCGGCAACAAGCTCAAG 358

Search completed: June 15, 2003, 03:41:52  
Job time : 1123.4 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 1662.94 Seconds  
(without alignments)  
11183.052 Million cell updates/sec

Title: US-09-513-888C-1\_COPY\_4912\_5550  
Perfect score: 639  
Sequence: 1 gaggtgaacgcgaaggtag.....tggaggcagggtcttcgccg 639

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	639	100.0	5492	9 AF123659	AF123659 Homo sapi
2	639	100.0	9108	9 AF123653	AF123653 Homo sapi
3	626.4	98.0	173264	2 AC025853	AC025853 Homo sapi
4	445	69.6	227884	2 AC114995	AC114995 Mus muscu
C 5	445	69.6	263546	2 AC099416	AC099416 Mus muscu
C 6	438.6	68.6	191210	2 AC108987	AC108987 Rattus no
7	434	67.9	1722	9 AF123658	AF123658 Homo sapi
8	374	58.5	1692	9 AF123657	AF123657 Homo sapi
C 9	232.4	36.4	8530	2 AC125907	AC125907 Rattus no
10	224.4	35.1	1614	9 AF123655	AF123655 Homo sapi
11	175.4	27.4	1515	9 AF123656	AF123656 Homo sapi
12	148.2	23.2	1612	9 BC005855	BC005855 Homo sapi
13	148.2	23.2	2275	9 HSM805394	BC006212 Homo sapi
14	148.2	23.2	2766	9 BC006212	BC006212 Homo sapi
15	148.2	23.2	5733	9 AB058716	AB058716 Homo sapi
16	148.2	23.2	181086	9 AL133215	AL133215 Human DNA
17	143.6	22.5	1935	9 AY029201	AY029201 Homo sapi
18	140.4	22.0	2099	9 AB046013	AB046013 Macaca fa
19	136.2	21.3	2767	10 BC014695	BC014695 Mus muscu
20	125.2	19.6	5257	9 AB011124	AB011124 Homo sapi
C 21	125.2	19.6	125856	9 HS1187M17	AL121891 Human DNA
C 22	123.4	19.3	103729	2 AC096620_3	Continuation (4 of
23	123.4	19.3	270700	2 AC105485	AC105485 Rattus no
C 24	120.2	18.8	168210	2 AC116700	AC116700 Mus muscu
C 25	118.2	18.5	167636	2 AC107097	AC107097 Rattus no
C 26	118.2	18.5	225045	2 AL833803	AL833803 Mus muscu
C 27	116.8	18.3	184865	2 AL807824	AL807824 Mus muscu
C 28	116.8	18.3	207683	2 AC098712	AC098712 Mus muscu
C 29	116	18.2	195673	2 AL807807	AL807807 Mus muscu
C 30	115.2	18.0	287323	2 AC093363	AC093363 Mus muscu
31	114.6	17.9	218760	2 AC105989	AC105989 Mus muscu
32	114.4	17.9	1150	14 HS4ULIR3	J02079 Epstein-Bar
33	114.4	17.9	1926	6 AX107940	AX107940 Sequence
34	114.4	17.9	2580	6 AR108994	AR108994 Sequence
C 35	114.4	17.9	5452	6 AR083151	AR083151 Sequence
36	114.4	17.9	5452	12 U02454	U02454 Cloning vec
37	114.4	17.9	9600	6 A92665	A92665 Sequence.1
38	114.4	17.9	9600	6 AR158345	AR158345 Sequence
39	114.4	17.9	10596	6 I25041	I25041 Sequence 15
40	114.4	17.9	10596	6 I30503	I30503 Sequence 15
41	114.4	17.9	10737	12 XXU02428	U02428 Cloning vec
42	114.4	17.9	10850	12 U02455	U02455 Cloning vec
43	114.4	17.9	172281	14 EBV	V01555 Epstein-Bar
44	114.4	17.9	184113	14 HS4B958RAJ	M80517 Epstein-Bar
C 45	114.2	17.9	155662	2 AC112388	AC112388 Rattus no

ALIGNMENTS

RESULT 1  
AF123659  
LOCUS AF123659 Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.  
DEFINITION AF123659 Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.  
ACCESSION AF123659  
VERSION AF123659.1 GI:4572475  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,

and its expression is altered in multiple human tumors  
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
99199287  
PUBMED 10097140  
REFERENCE 2 (bases 1 to 5492)  
AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,  
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
Direct Submission  
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
JOURNAL Institute, 233S 10th street, Philadelphia, PA 19107, USA  
FEATURES  
Location/Qualifiers  
1..5492  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="8p22"  
1..5492  
/gene="FEZ1"  
112..1902  
/gene="FEZ1"  
/codon\_start=1  
/product="FEZ1"  
/protein\_id="AAD23840.1"  
/db\_xref="GI:4572476"  
/translation="MGSVSSLLISGHSHKCRASQYKLRKSHLKLNRYSDDLRF  
PFLMPFNSQLEMGSEKAVPTAFKVLPRSGAILHSSPESASHQHPAPDPKPEQ  
ELKPLGSLGSDSRNSMSLTHSTSSYQLDPLVTPVGTFRFGSAHNTQGV  
LQDNMMSLKALSFSDGSKLHGNKADKSPSVRSTDECSTIQLEQLKLRREGA  
LQKLRPEEKELASSLAYERPRRCDELEGPKNKLOKASOKSQAQOVLHLO  
VLQOQKRLQRLQESLMKQDLETLKLSYREKTSFGPALBETOWEVCQKSGEIS  
ENLQKKNKNEALREKYNLLEQLRAQALARDMPPTFFEDVPALORELELR  
AELREKQHDQMSGFQHERLVWKEEKVITQYQKLOQSVAMYQNRQLEKALQ  
LARGSAGEPLEVDLEGADIPYEDIIATEI"  
BASE COUNT 1137 a 1704 c 1565 g 1086 t  
ORIGIN  
Query Match 100.0%; Score 639; DB 9; Length 5492;  
Best Local Similarity 100.0%; Pred. No. 7e-78;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAGGTGAACGCAAGCTAGCAGATCTGGTCTCAAGGCACACGCTGAAGACACGCGG 60  
DB 1321 GAGGTGAACGCAAGCTAGCAGATCTGGTCTCAAGGCACACGCTGAAGACACGCGG 1380  
QY 61 GCGAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 120  
DB 1381 GCGAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 1440  
QY 121 GGCCTGGAGCTGGAGCTGTGTAGAAATGAGCTGCAGCGCAAGAAACGAGCGGAGCTG 180  
DB 1441 GGCCTGGAGCTGGAGCTGTGTAGAAATGAGCTGCAGCGCAAGAAACGAGCGGAGCTG 1500  
QY 181 CTGGCGGAGAAAGTGAACCTGTGAGAGCAGGAGCTGCAGGAGCTGGGGCCCGAGCCGCC 240  
DB 1501 CTGGCGGAGAAAGTGAACCTGTGAGAGCAGGAGCTGCAGGAGCTGGGGCCCGAGCCGCC 1560  
QY 241 CTGGCCCGCGACATGGGGCGGCCCGACCTTCCCGAGGAGCTGCCCTGCCCTGACGCGGAG 300  
DB 1561 CTGGCCCGCGACATGGGGCGGCCCGACCTTCCCGAGGAGCTGCCCTGCCCTGACGCGGAG 1620  
QY 301 CTGGAGCGGCTGGGGCCCGAGCTGGGAGGAGCGCGCAAGGCATGACAGATGTCCTCG 360  
DB 1621 CTGGAGCGGCTGGGGCCCGAGCTGGGAGGAGCGCGCAAGGCATGACAGATGTCCTCG 1680  
QY 361 GCCTTCCAGCATGAGCGGCTCGTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
DB 1681 GCCTTCCAGCATGAGCGGCTCGTGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740  
QY 421 AAACAGCTCAGCAGAGCTACGTGGCCATGTACACGCGGAACAGCGCTGGAGAGGCC 480  
DB 1741 AAACAGCTCAGCAGAGCTACGTGGCCATGTACACGCGGAACAGCGCTGGAGAGGCC 1800

QY 481 CTGCAGCAGCTGGCAGCTGGGGACAGCGCGGGAGCCCTTGGAGTTGACCTGGAAGGG 540  
DB 1801 CTGCAGCAGCTGGCAGCTGGGGACAGCGCGGGAGCCCTTGGAGTTGACCTGGAAGGG 1860  
QY 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGTGCTCTGGGAAGGC 600  
DB 1861 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGTGCTCTGGGAAGGC 1920  
QY 601 GAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 639  
DB 1921 GAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 1959  
RESULT 2  
AF123653 9108 bp DNA linear PRI 07-APR-1999  
LOCUS Homo sapiens FEZ1 (FEZ1) gene, complete cds.  
DEFINITION  
ACCESSION AF123653  
VERSION AF123653.1 GI:4572463  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 9108)  
AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,  
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
TITLE The FEZ1 gene at chromosome 9p22 encodes a leucine-zipper protein,  
and its expression is altered in multiple human tumors  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
MEDLINE 99199287  
PUBMED 10097140  
REFERENCE 2 (bases 1 to 9108)  
AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H.,  
Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
Direct Submission  
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
JOURNAL Institute, 233S 10th street, Philadelphia, PA 19107, USA  
FEATURES  
Location/Qualifiers  
1..9108  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
/gene="FEZ1"  
/gene="FEZ1"  
/gene="FEZ1"  
/product="FEZ1"  
/protein\_id="AAD23834.1"  
/db\_xref="GI:4572464"  
/translation="MGSVSSLLISGHSHKCRASQYKLRKSHLKLNRYSDDLRF  
PFLMPFNSQLEMGSEKAVPTAFKVLPRSGAILHSSPESASHQHPAPDPKPEQ  
ELKPLGSLGSDSRNSMSLTHSTSSYQLDPLVTPVGTFRFGSAHNTQGV  
LQDNMMSLKALSFSDGSKLHGNKADKSPSVRSTDECSTIQLEQLKLRREGA  
LQKLRPEEKELASSLAYERPRRCDELEGPKNKLOKASOKSQAQOVLHLO  
VLQOQKRLQRLQESLMKQDLETLKLSYREKTSFGPALBETOWEVCQKSGEIS  
ENLQKKNKNEALREKYNLLEQLRAQALARDMPPTFFEDVPALORELELR  
AELREKQHDQMSGFQHERLVWKEEKVITQYQKLOQSVAMYQNRQLEKALQ  
LARGSAGEPLEVDLEGADIPYEDIIATEI"  
BASE COUNT 2021 a 2628 c 2504 g 1952 t  
ORIGIN

Query Match 100.0%; Score 639; DB 9; Length 9108;  
Best Local Similarity 100.0%; Pred. No. 6.3e-78;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 GAGGTGAACGCGGCTAGCGAGATCTCTGGGTCTCAAGGCACACGCTGAAGGACACGCGG 60  
 Db 4972 GAGGTGAACGCGGCTAGCGAGATCTCTGGGTCTCAAGGCACACGCTGAAGGACACGCGG 5031  
 QY 61 GCGAGCTGGAGGGCTGAGCTAGGACCCAGGACCTGGAGGGGCGCCCTGGCGACCAAG 120  
 Db 5032 GCGAGCTGGAGGGCTGAGCTAGGACCCAGGACCTGGAGGGGCGCCCTGGCGACCAAG 5091  
 QY 121 GCGCTGAGCTGGAGGCTCTGAGAAATGAGCTGCGAGCGCAAGAAAGCGAGCGGAGCTG 180  
 Db 5092 GCGCTGAGCTGGAGGCTCTGAGAAATGAGCTGCGAGCGCAAGAAAGCGAGCGGAGCTG 5151  
 QY 181 CTGCGGAGAAAGTGAACCTGCTGAGCAGGAGCTGCGAGGAGCTCGCGGCGCCAGCCGCC 240  
 Db 5152 CTGCGGAGAAAGTGAACCTGCTGAGCAGGAGCTGCGAGGAGCTCGCGGCGCCAGCCGCC 5211  
 QY 241 CTGCGGCGCGCATGCGGCGCGCCACCTTCCCGAGGAGGCTCCCTGCGCTGCGAGCGGAG 300  
 Db 5212 CTGCGGCGCGCATGCGGCGCGCCACCTTCCCGAGGAGGCTCCCTGCGCTGCGAGCGGAG 5271  
 QY 301 CTGAGCGGCTGCGGCGCGGAGCTGCGGAGGAGCGGCGGAGGAGGAGGAGGAGGAGGAG 360  
 Db 5272 CTGAGCGGCTGCGGCGCGGAGCTGCGGAGGAGCGGCGGAGGAGGAGGAGGAGGAGGAG 5331  
 QY 361 GCGTTCAGCATGACGCGCTCGTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
 Db 5332 GCGTTCAGCATGACGCGCTCGTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5391  
 QY 421 AAACAGCTGCGACAGCTAGCTGCGCATGTACCGAGCGGAGGAGGAGGAGGAGGAGGAGG 480  
 Db 5392 AAACAGCTGCGACAGCTAGCTGCGCATGTACCGAGCGGAGGAGGAGGAGGAGGAGGAGG 5451  
 QY 481 CTGACAGCTGCGACAGCTAGCTGCGCATGTACCGAGCGGAGGAGGAGGAGGAGGAGGAGG 540  
 Db 5452 CTGACAGCTGCGACAGCTAGCTGCGCATGTACCGAGCGGAGGAGGAGGAGGAGGAGGAGG 5511  
 QY 541 GCTGACATCCCTACGAGGACATCATAGCACTGAGATCTGAGGGGCTGCGCTGGGAAGGC 600  
 Db 5512 GCTGACATCCCTACGAGGACATCATAGCACTGAGATCTGAGGGGCTGCGCTGGGAAGGC 5571  
 QY 601 GAGCTGGGACCTGGACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 639  
 Db 5572 GAGCTGGGACCTGGACTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5610

RESULT 3  
 AC025853  
 LOCUS Homo sapiens chromosome 8 clone RP11-353K12 map 8, HTG 17-JUN-2002  
 DEFINITION IN PROGRESS \*\*\*, 3 unordered pieces.  
 ACCESSION AC025853  
 VERSION AC025853  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 173264)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Brown,A., Borkett,G.,  
 Anderson,S., Baldin,J., Bouckhelter,B., Brown,A., Bastien,V., Beda,F.,  
 Boguslavsky,L., Bouckhelter,B., Brown,A., Bastien,V., Beda,F.,  
 Campopiano,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,  
 Collimore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Grand-Pierre,N., Grant,G., Hagos,B., Headford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,

Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,  
 Melcham,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Pierre,N.,  
 Pibani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 173264)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
 Bouckhelter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collimore,A.,  
 Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,  
 Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,  
 Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,  
 Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,  
 Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,  
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
 Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,  
 Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,  
 Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,  
 Stojanovic,N., Subramanian,A., Talamas,J., Tassiliev,H.,  
 Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,  
 Vasilev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jun 17, 2002 this sequence version replaced gi:21321864.  
 All repeats were identified using RepeatMasker:  
 Smit,A.P.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 -----Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 -----Project Information  
 Center project name: L7454  
 Center clone name: 353\_K\_12  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 39461: contig of 39461 bp in length  
 \* 39462 39561: gap of 100 bp  
 \* 39562 108347: contig of 68786 bp in length  
 \* 108348 108447: gap of 100 bp  
 \* 108448 173264: contig of 64817 bp in length.  
 Location/Qualifiers  
 1. 173264  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8"  
 /clone="RP11-353K12"  
 FEATURES  
 source

```

/clone.lib="RPC1-11 Human Male BAC"
BASE COUNT 47355 a 38864 c 39235 g 47493 t 317 others
ORIGIN
Query Match 98.0%; Score 626.4; DB 2; Length 173264;
Best Local Similarity 99.7%; Pred. No. 1.7e-76;
Matches 638; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GAGGTGAACGCCAAGCTAGCAGATCTCGGTCTCAAGGCACACTGAAGACACGGG 60
DB 27212 GAGGTGAACGCCAAGCTAGCAGATCTCGGTCTCAAGGCACACTGAAGACACGGG 27271
QY 61 GCACAGCTGGAGGGCTCGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 120
DB 27272 GCACAGCTGGAGGGCTCGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 27331
QY 121 GGCCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAACGAGCGCGGAGCTG 180
DB 27332 GGCCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAAACGAGCGCGGAGCTG 27391
QY 181 CTCGGGAGAGGTGAACCTGCTGAGCAGAGCTGCGAGGCTCGGGGCCAGCCGCC 240
DB 27392 CTCGGGAGAGGTGAACCTGCTGAGCAGAGCTGCGAGGCTCGGGGCCAGCCGCC 27451
QY 241 CTGGCCCGGACATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCGCGGAGGAG 300
DB 27452 CTGGCCCGGACATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCGCGGAGGAG 27511
QY 301 CTGGAGCGCTGCGGGCGGAGCTGCGGGAGGCGGCAAGGCATGACAGATGTCCTCG 360
DB 27512 CTGGAGCGCTGCGGGCGGAGCTGCGGGAGGCGGCAAGGCATGACAGATGTCCTCG 27571
QY 361 GCTTCCAGCATGAGCGGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 27572 GCTTCCAGCATGAGCGGCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 27631
QY 421 AAACAGCTGCAGCAGCTACGTGGCCATGTACACGCGGAAACGAGCGCTGGAGAAGCC 480
DB 27632 AAACAGCTGCAGCAGCTACGTGGCCATGTACACGCGGAAACGAGCGCTGGAGAAGCC 27691
QY 481 CTGACAGCTGGCAGCTGGGACAGCGCGGGAGCGCCCTTGG-AGGTGACCTGGAAGG 539
DB 27692 CTGACAGCTGGCAGCTGGGACAGCGCGGGAGCGCCCTTGG-AGGTGACCTGGAAGG 27751
QY 540 GCTCAGATCCCTACGAGGACATCATAGCAGCTGAGTCTGAGGGCTGCTGGAGG 599
DB 27752 GCTCAGATCCCTACGAGGACATCATAGCAGCTGAGTCTGAGGGCTGCTGGAGG 27811
QY 600 CGAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 639
DB 27812 CGAGTCTGGGACCTGGCACTGGGAGGAGGAGGCTCTCCCG 27851

RESULT 4
AC114995 227884 bp DNA linear HTG 26-JUN-2002
LOCUS Mus musculus clone RP23-171P5, WORKING DRAFT SEQUENCE, 5 ordered
DEFINITION pieces.
ACCESSION AC114995
VERSION AC114995.3 GI:21592059
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 227884)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gordon,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Lander,E., Lehotzky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
Minova,C., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olivier,J., Peterson,K., Phunkhang,P., Rieback,M., Riley,R., Rise,C.,
Raymond,C., Retta,R., Rieback,M., Rieback,M., Rieback,M., Rieback,M.,
Roman,J., Roetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodores,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 227884)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
TITLE
JOURNAL
COMMENT

```

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 5 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 301: contig of 301 bp in length  
 \* 302 401: gap of 100 bp  
 \* 402 5831: contig of 5430 bp in length  
 \* 5832 5931: gap of 100 bp  
 \* 5932 173029: contig of 167098 bp in length  
 \* 173030 173129: gap of 100 bp  
 \* 173130 210061: contig of 36932 bp in length  
 \* 210062 210161: gap of 100 bp  
 \* 210162 227884: contig of 17723 bp in length.

## FEATURES

```

source
1..227884
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-171Ps"
/clone.lib="RPI-23 Female Mouse BAC"

misc_feature
1..301
/note="assembly_fragment"
clone_end:SP6
vector_side:left"

misc_feature
402..5831
/note="assembly_fragment"

misc_feature
5932..173029
/note="assembly_fragment"

misc_feature
173130..210061
/note="assembly_fragment"

misc_feature
210162..227884
/note="assembly_fragment"
clone_end:T7
vector_side:right"

BASE COUNT 65784 a 48829 c 49902 g 62967 t 402 others
ORIGIN

Query Match 69.6%; Score 445; DB 2; Length 227884;
Best Local Similarity 82.2%; Pred. No. 6.4e-52;
Matches 530; Conservative 0; Mismatches 100; Indels 15; Gaps 1;

QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACGGG 60
|||||
DB 87422 GAGGTGAACACTAAGCCAGCAGATCTCTAGTCTGAAGGCACAGCTGAAGATACCCGG 87481

QY 61 GCCAAGCTGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCCCTGCGCCACCAAG 120
|||||
DB 87482 GGCAGCTGGATGAATGAGCTGAAGACACAGGACTTGGAGTGGCTGCGCACCAAA 87541

QY 121 GGCCTGGAGCTGAGGTCTGTAGATGATGCTGCGCGGCAAGAACGAGCGGAGCTG 180
|||||
DB 87542 GGGCTGGAGCTGAGGTCTGTAGAACCGAATCTGCGCGGCAAGAACGAGCAGACTG 87601

QY 181 CTGCGGGAAGGTGAACCTGCTGGAGCAGGAGCTGCGGCGCCAGGCGG-- 238
|||||
DB 87602 CTGAGAGAGAAGGTGNACCTTCTAGACAGAGATTGATGAGACTAGGGCCCGAGCTGCT 87661

QY 239 -----CCTGCGCCCGCAGATG7GGGGCCGCCACCTTCCCGAGGACGTCCT 285
|||||
DB 87662 CTGCATCCGGCACCCCTCGGGGCTCCAGGGGTGGGCTCACTTTCTCAGAGGACATCCCT 87721

QY 286 GGCCTGCAGCGGAGCTGGAGGGCTGCGGGCGGAGCTCGGGAGGAGCGGCAAGCCAT 345
|||||
DB 87722 GCTCTGACGGGAGCTGGATCGGCTGCGGGCGGAGCTGAAGAGGAGGAGCAAGGCCAT 87781

QY 346 GACCAGATCTCTCGGGCTTCCAGCATGATGCGGCTGCTGTGAAGAGGAGGAGGAGAG 405
|||||
DB 87782 GACCAGATCTCTCGGGATCCAGCATGAGCCGGTGGTGTGAAGAGGAGGAGAGAAA 87841

QY 406 GTGATTTCAGTACAGAAAACAGCTGCAGCAGAGACTAGTGGCCATGTATCCAGCGGAACCA 465

```

```

Db 87842 GTGATCCAGTACCAAGGCAGCTGCAGCAGAGCTACTTTGGCCATGTACCAAGCAACCA 87901
|||||
QY 466 CGCTGGAGAGAGCCCTCCAGCAGCTGCAGCTGGGGACAGCGCGGGAGCCCTTTGGAG 525
|||||
DB 87902 GCTCTAGAGAGAGCCCTCCAGCAGCTGCAGCTGGCCGTGGGGATGTTCCAGGGGAACCCCTTGA 87961
|||||
QY 526 GTTGACCTGGAGGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGG 585
|||||
DB 87962 ATTGACCTGGAGGGAGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGG 88021
|||||
QY 586 GCTGCTGGAGAGGGAGCTCTGGGACCTGGCACTGGGAGGGAGG 630
|||||
DB 88022 GCTGCTAGGGGTGGAGGGCTGGGACTGGACCCAGCAAGAGCAGG 88066
|||||

RESULT 5
AC099416/c
LOCUS
DEFINITION AC099416 263546 bp DNA linear HTG 23-MAY-2002
SEQUENCE, 8 unordered pieces.
ACCESSION AC099416
VERSION AC099416.2 GI:21105058
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 263546)
AUTHORS McPherson,J.D. and Waterston,R.H.
JOURNAL The sequence of Mus musculus clone
Unpublished
REFERENCE
2 (bases 1 to 263546)
AUTHORS McPherson,J.D. and Waterston,R.H.
JOURNAL Direct Submission
Submitted (14-NOV-2001) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE
3 (bases 1 to 263546)
AUTHORS McPherson,J.D. and Waterston,R.H.
JOURNAL Direct Submission
Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On May 23, 2002 this sequence version replaced gi:16924178.
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0122M11
----- Summary Statistics -----
Sequencing vector: M13; 32x
Sequencing vector: plasmid; 68x
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.950319
Consensus quality: 278310 bases at least Q40
Consensus quality: 282938 bases at least Q30
Consensus quality: 286149 bases at least Q20
Insert size: 250000; agarose-fp
Insert size: 387570; sum-of-contigs
Quality coverage: 25.74 in Q20 bases; agarose-fp
Quality coverage: 16.86 in Q20 bases; sum-of-contigs

```

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will

```

* be preserved.
1 1204: contig of 1204 bp in length
1205 1304: gap of unknown length
1305 2524: contig of 1220 bp in length
2525 2624: gap of unknown length
2625 4136: contig of 1512 bp in length
4137 4236: gap of unknown length
4237 16434: contig of 12198 bp in length
16435 16534: gap of unknown length
16535 32106: contig of 15572 bp in length
32107 32206: gap of unknown length
32207 140224: contig of 108018 bp in length
140225 140324: gap of unknown length
140325 262933: contig of 122369 bp in length
262934 262794: gap of unknown length
262794 263546: contig of 753 bp in length.
FEATURES
    Location/Qualifiers
        1..263546
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /chromosomes="UNK"
            /clone="RP23-122M11"
        misc_feature
            1..1204
                /note="assembly_name:Contig127"
        misc_feature
            1305..2524
                /note="assembly_name:Contig157"
        misc_feature
            2625..4136
                /note="assembly_name:Contig360"
        misc_feature
            4237..16434
                /note="assembly_name:Contig396"
        misc_feature
            16535..32106
                /note="assembly_name:Contig397"
        misc_feature
            32207..140224
                /note="assembly_name:Contig398"
        misc_feature
            140325..262693
                /note="assembly_name:Contig399"
        misc_feature
            262794..263546
                /note="assembly_name:Contig154"
BASE COUNT 76628 a 56149 c 55623 g 74436 t 710 others
ORIGIN
Query Match 69.6%; Score 445; DB 2; Length 263546;
Best Local Similarity 82.2%; Pred. No. 6.2e-52;
Matches 530; Conservative 0; Mismatches 100; Indels 15; Gaps 1;
QY 1 GAGGTGAACCGAAGCTAGCAGATCTCTGAGTCTCAAGGCAAGCTGAAGCAGCGGG 60
Db 247464 GAGGTGAACCACTAAGCCAGCGAGATCTCTGAGGCAAGCTGAAGGATACCGGG 247405
QY 61 GCGAGCTGAGGGCTTGAGCTGAGGACCGAGGCTGGAGGCGCTGCGCACCAG 120
Db 247404 GCGAGCTGAGTGAATGAGCTGAAGACAGGACTTGGAGTGCCCTGCGCACCAG 247345
QY 121 GCGCTGAGCTGAGTCTGTGAGATGAGCTGCGCAAGCAAGACGAGCGGAGCTG 180
Db 247344 GCGCTGAGCTGAGTCTGTGAGACGAACTGCGCGCAAGAGATGAGGACGACTG 247285
QY 181 CTGCGGGAAGGTGAACCTGCTGAGCAGAGAGCTGCGAGGCTGGGGCCCGCGCG-- 238
Db 247284 CTGAGAGAAGGTGAACCTTCTAGAACAGGAGTTGATGAGTACGCGGCCCGAGCTGT 247225
QY 239 -----CCTGCGCCCGAGACATGGGGCGCCGACCTTCCCGAGGACGTCCT 285
Db 247224 CTGCATCCGGCACCCTCGGGGCTCCAGGGGTTGGGCTCCTTTCTCAGAGACATCCCT 247165
QY 286 GCGCTGAGCGGAGCTGAGCGGGCTGCGGGCGGAGCTCGGGAGAGCGGCAAGGCCAT 345
Db 247164 GCTCTGACGGGAGCTGATCGGCTGCGGGCGGAGCTGAAGGAGGAGGCAAGGCCAT 247105
QY 346 GACCAGATGCTCGGGCTTCAGACATGAGCGGCTCGTGTGAAGAGGAGGAGAGAG 405
Db 247104 GACCAGATGCTCGGGATTCAGCATGAGCGGTTGTTGTGGAAGAGGAGAGAGAA 247045

```

406 GTGATTTCAGTACCAGAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAGCGAACACAG 465  
 247044 GTGATTCAGTACCAGAACAGCAGCTGCAGCAGAGCTACTTGGCCATGTACCAGCGAACACAG 246985  
 466 CGCCTGGAGAAAGGCCCTGCAGCAGCTGCAGCAGCTGGGGAACAGCGCGGGAGGCCCTTTGAG 525  
 246984 CGTCTAGAGAAGGCCCTGCAGCAGCTGCAGCAGCTGGCCGTGGGATGTCTCCAGGGGACCCCTTTCAA 246925  
 526 GTTGACCTGGAAAGGGCTGACATCCCTACAGAGACATCATAGCAGCTGAGATCTGAGGG 585  
 246924 ATTGACCTGGAGGAGCTGACATCCCTACAGAGACATCATAGCAGCTGAGATCTGAGGG 246865  
 586 GCTGCTGGGAAGCGAGCTCTGGGACCTGCGACCTGGAGCGAGG 630  
 246864 CTGCTAGGCTGGAGGCTGGGACCTGGGACCTGGACCCAGAAAGCAGG 246820

RESULT 6  
 AC108987/c 191210 bp DNA linear HTG 13-JUL-2002  
 LOCUS Rattus norvegicus clone CH230-115K1, \*\*\* SEQUENCING IN PROGRESS  
 DEFINITION \*\*\* 78 unordered pieces.  
 ACCESSION AC108987  
 VERSION AC108987.3 GI:21737647  
 KEYWORDS HTG; HTGS PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1 (bases 1 to 191210)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barbakia,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhatz,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Caron,T.P., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homsai,P., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okuwonu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
 Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Umani,K., Vasquez,L., Veta,V., Villalón,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.  
 Direct Submission  
 TITLE  
 JOURNAL  
 Unpublished

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 191210)

Worley, K.C.

Direct Submission

Submitted (03-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 191210)

Worley, K.C.

Direct Submission

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On July 12, 2002 this sequence version replaced gi:18846600.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: GPMW

Center clone name: CH230-115K1

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 115719 bases at least Q40

Consensus quality: 122885 bases at least Q30

Consensus quality: 129025 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 78 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1041: contig of 1041 bp in length

\* 1042 1141: gap of unknown length

\* 1142 2262: contig of 1121 bp in length

\* 2263 2362: gap of unknown length

\* 2363 3338: contig of 1176 bp in length

\* 3339 3638: gap of unknown length

\* 3639 4775: contig of 1137 bp in length

\* 4776 4875: gap of unknown length

\* 4876 6050: contig of 1175 bp in length

\* 6051 6151: gap of unknown length

\* 6151 7760: contig of 1510 bp in length

\* 7761 9331: contig of 1571 bp in length

\* 9332 9431: gap of unknown length

\* 9432 11027: contig of 1596 bp in length

\* 11028 11127: gap of unknown length

\* 11128 12769: contig of 1642 bp in length

\* 12770 12869: gap of unknown length

\* 12870 14092: contig of 1223 bp in length

\* 14093 14192: gap of unknown length

\* 14193 15228: contig of 1036 bp in length

\* 15229 15328: gap of unknown length

\* 15329 16709: contig of 1381 bp in length

\* 16710 16809: gap of unknown length

\* 16810 18195: contig of 1386 bp in length

\* 18196 18295: gap of unknown length

\* 18296 19412: contig of 1117 bp in length

\* 19413 19512: gap of unknown length

\* 19513 20959: contig of 1447 bp in length

\* 20960 21059: gap of unknown length

\* 21060 22161: contig of 1102 bp in length

\* 22162 22661: gap of unknown length

\* 22662 23587: contig of 1326 bp in length

\* 23588 23687: gap of unknown length

\* 23688

23688 25231: contig of 1544 bp in length

25232 25331: gap of unknown length

25332 27135: contig of 1804 bp in length

27136 27235: gap of unknown length

27236 29176: contig of 1940 bp in length

29177 29275: gap of unknown length

29276 30846: contig of 1571 bp in length

30847 30946: gap of unknown length

30947 32090: contig of 1144 bp in length

32091 32190: gap of unknown length

32191 34039: contig of 1849 bp in length

34040 34139: gap of unknown length

34140 35408: contig of 1269 bp in length

35409 35508: gap of unknown length

35509 37018: contig of 1510 bp in length

37019 37118: gap of unknown length

37119 38260: contig of 1142 bp in length

38261 38360: gap of unknown length

38361 40083: contig of 1723 bp in length

40084 40183: gap of unknown length

40184 41200: contig of 1017 bp in length

41201 41300: gap of unknown length

41301 43365: contig of 2065 bp in length

43366 43465: gap of unknown length

43467 45343: contig of 1878 bp in length

45344 45443: gap of unknown length

45444 46987: contig of 1544 bp in length

46988 47087: gap of unknown length

47088 49404: contig of 2317 bp in length

49405 49504: gap of unknown length

49505 50823: contig of 1319 bp in length

50824 50923: gap of unknown length

50924 52451: contig of 1528 bp in length

52452 52551: gap of unknown length

52552 54441: contig of 1890 bp in length

54442 54541: gap of unknown length

54542 56442: contig of 1901 bp in length

56443 56542: gap of unknown length

56543 57749: contig of 1207 bp in length

57750 57849: gap of unknown length

57850 60138: contig of 2289 bp in length

60139 60238: gap of unknown length

60239 62467: contig of 2229 bp in length

62468 62567: gap of unknown length

62568 64364: contig of 1797 bp in length

64365 64464: gap of unknown length

64465 66889: contig of 2425 bp in length

66890 66989: gap of unknown length

66990 68830: contig of 1841 bp in length

68831 68930: gap of unknown length

68931 71226: contig of 2296 bp in length

71227 71326: gap of unknown length

71327 73712: contig of 2386 bp in length

73713 73812: gap of unknown length

73813 76871: contig of 3059 bp in length

76872 76971: gap of unknown length

76972 79844: contig of 2873 bp in length

79845 79944: gap of unknown length

79945 81259: contig of 1315 bp in length

81260 81359: gap of unknown length

81360 83690: contig of 2331 bp in length

83691 83790: gap of unknown length

83791 86195: contig of 2405 bp in length

86196 86295: gap of unknown length

86296 88065: contig of 1770 bp in length

88066 88165: gap of unknown length

88166 90384: contig of 2219 bp in length

90385 90484: gap of unknown length

90485 91999: contig of 1515 bp in length

92000 92099: gap of unknown length

92100 94583: contig of 2484 bp in length

94584 94683: gap of unknown length

94684 98297: contig of 3614 bp in length

Query Match 68.6%; Score 439.6; DB 2; Length 191210;  
Best Local Similarity 82.2%; Pred. No. 4.9e-51;  
Matches 537; Conservative 0; Mismatches 94; Indels 22; Gaps 2;

QY 1 GAGGTGAAGCCCAAGCTAGCAGATCTTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60  
DB 105989 GAGGTGAATCTTAGCCAGTGAATCTTACCTTAGAGGCACAGCTGAAGGACACCGG 105930

QY 61 GCACAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCCCTTGCACCAAG 120  
DB 105929 GCACAGCTGGAGGATGAGTTGAAGACACAGGACTTGGAGAGCCCTTGGCCACCAAG 105870

QY 121 GCCTTGGAGCTGAGGCTGTGAGATGAGCTGACGCGCAAGAACAGAGCGGAGCTG 180  
DB 105869 GGGCTAGAGCTGGAGGCTGTGAGATGAATGAACTGCAGCGCAAGAAAGATGAGGACAGCTG 105810

QY 181 CTGCGGGAGAGTGAACCTGCTGAGCAGGAGCTGCAGGAGCTGCGGGCCCGAGCCGCC 240  
DB 105809 CTGAGAGAGAGTGAACCTTTAGAACAGGAGCTGCTGGAGCTGCGGGCCCGAGCCGCC 105750

QY 241 CTGCGCCCGACATCGGGCC-----GCCACCTTCCCGAGGAC 279  
DB 105749 CTGCATCGAGATGCAGCACCCCTGGGCCCTCCAGGGATTGGGCTGACTTTCTCAGAGGAC 105690

QY 280 GTCCCTGCCCTCAGCGGAGCTGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCA 339  
DB 105689 ATCCCTGTCTCAGCGGAGCTGAGCTCGGCTCGCGGCTGAGCTGAAGAGGAGAGGCA 105630

QY 340 GCCCATGACCATGCTCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAG 399  
DB 105629 GCCCATGACCATGCTCTCGGGATTCCAGCATGAGCGGCTTGGTGTGGAAGGAGAGAG 105570

QY 400 GAGAAGGTGATTCAGTACCAAGAACAGCTGCAGCAGAGCTACGTGGCCACTGTACCAAGCG 459  
DB 105569 GAGAAGGTGATTCAGTACCAAGAACAGCTGCACAGAGCTACTTGGCCACTGTACCAAGCG 105510

QY 460 AACACAGCGCTGAGAGCGCTGACAGCTGACAGCTGCGGAGAGCGCGGAGCGCC 519  
DB 105509 AACACAGCGCTGAGAGCGCTGACAGCTGCGGAGAGCTGCGGAGATTTGCAAGGGAGCGCC 105450

QY 520 TTGAGGCTTACCTGGAAGGGCTGACATCCCTACGAGGAGCATCATAGCCACTGAGATC 579  
DB 105449 TTGAAATTGACTGGAGGAGCTGACATTCCTATGAGGATATCATAGCCACTGAGATC 105390

QY 580 TGAGGGCTGCTT-GGGAAGCGAGTCTGGGAGCTTGGGAGCTTGGAGCGAGCG 631  
DB 105389 TGAGGGCTTACCTCGGGAGTGGAGGCTGGGGAGCTAGACCCAGAAAGCAGGG 105337

RESULT 7  
AF123658  
LOCUS  
DEFINITION Homo sapiens clone G3612 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999 complete cds.  
ACCESSION AF123658  
VERSION AF123658.1 GI:4572473  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1722)  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
MEDLINE 99199287  
PUBMED 10097140  
REFERENCE 2 (bases 1 to 1722)  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H., Mori,M., Fidanza,V., Alder,H. and Croce,C.M.

TITLE Direct Submission  
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 233S 10th street, Philadelphia, PA 19107, USA  
FEATURES  
source  
1. 1722  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
/clone="G3612"  
1. 1722  
/gene="FEZ1"  
1. 1722  
/gene="FEZ1"  
/note="alternatively spliced"  
/codon\_start=1  
/product="FEZ1"  
/protein\_id="A457247.1"  
/db\_xref="GI:4572474"  
/translation="MGVSLSISGHSFKHCRASQYKLRKSSHLKLNRYSDGLLRF  
GFSDSGHSGSSXMGKSEDFYIKVQKARGSHHPDYLTSSDGLGQAGVDFDPST  
PKLMPFSNOLMGSEKGAFTAPKPLPRSGALHSSPESASHLPAPDKPKQ  
ELKPLCSGALSLSGRNMSLPTHTSSSYQLDPLVTPVPTSRFGSSAHNITGIV  
LQDNMMLKALSFSDGSKLGHNKADKGPSCVRSPTSTDECSIQLEOKLBRGA  
LQKORSFEKELEASSLAYERPRRDELEGPCKGNKQKQASQSQRAQQVHLQ  
VLQQRQRQRQLESLMKEDLLETKLRSYERKTSFGPALEETQWEVCQSGEIS  
LLQQLQKESQTEVNAKASEILGLKQKLDTRKLEGLRDTQDLEALRTKGLLEVC  
ENELQRKNEALLERKVNLEQLQELRAQAALARDMPPTTFEDVPALQRELERLV  
WKEKEKVIQYQKQLQSYVAMYQNRQLEKALQQLARGDSAGEPLEVLEADIPYE  
DIATEI"  
BASE COUNT 381 a 541 c 563 g 237 t  
ORIGIN  
Query Match 67.9%; Score 434; DB 9; Length 1722;  
Best Local Similarity 88.1%; Pred. No. 5.5e-50;  
Matches 513; Conservative 0; Mismatches 0; Indels 69; Gaps 1;  
QY 1 GAGGTGAAGCCCAAGCTAGCAGATCTTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60  
DB 1210 GAGGTGAAGCCCAAGCTAGCAGATCTTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 1269

QY 61 GCACAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 120  
DB 1270 GCACAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 1329

QY 121 GCCTTGGAGCTGGAGGCTGTGAGAATGAGCTGAGCGCAAGAAAGAACGAGCGGAGCTG 180  
DB 1330 GCCTTGGAGCTGGAGGCTGTGAGAATGAGCTGAGCGCAAGAAAGAACGAGCGGAGCTG 1389

QY 181 CTGCGGGAGAGGTGAACCTCTGAGCAGGAGCTGAGGAGCTGCGGGCCCGAGCCGCC 240  
DB 1390 CTGCGGGAGAGGTGAACCTCTGAGCAGGAGCTGAGGAGCTGCGGGCCCGAGCCGCC 1449

QY 241 CTGCGCCCGACATGGGGCCGCCACCTTCCCAGAGGAGCTCCCTGCGCTGAGCGGAG 300  
DB 1450 CTGCGCCCGACATGGGGCCGCCACCTTCCCAGAGGAGCTCCCTGCGCTGAGCGGAG 1509

QY 301 CTGAGAGCGGCTGCGGGCCCGAGCTGCGGGAGGAGCGGCAAGGCCCATGACCATGTCTCTCG 360  
DB 1510 CTG-----GAGCGGCTCGTGTGGAAGGAGGAGAGAGAGAGGTGATTCAGTACCAAG 1512

QY 361 GCCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGAGGTGATTCAGTACCAAG 420  
DB 1513 -----GAGCGGCTCGTGTGGAAGGAGGAGAGAGAGAGGTGATTCAGTACCAAG 1560

QY 421 AAACAGCTGACGAGAGCTAGTGGCCATGTATACAGCGGAAACAGCGCTGAGAGAGGCC 480  
DB 1561 AAACAGCTGACGAGAGCTAGTGGCCATGTATACAGCGGAAACAGCGCTGAGAGAGGCC 1620

QY 481 CTGAGCAGCTGGCACTGTGGGAGCAGCGCCCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 540  
DB 1621 CTGAGCAGCTGGCACTGTGGGAGCAGCGCCCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 1680

```

QY 541 GCTGATCCCTTACGAGACATCATAGCCACTGAGATCTGA 582
|||||
Db 1681 GCTGATCCCTTACGAGACATCATAGCCACTGAGATCTGA 1722
|||||

RESULT 8
AF123657
LOCUS Homo sapiens clone G3611 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999
DEFINITION complete cds.
ACCESSION AF123657
VERSION AF123657.1 GI:4572471
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1692)
REFERENCE
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 99199287
PUBMED 10097140
REFERENCE
AUTHORS 2 (bases 1 to 1692)
Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
Direct Submission
TITLE Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
JOURNAL Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES
source
1..1692
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="G3611"
1..1692
/gene="FEZ1"
1..1692
/gene="FEZ1"
/translation="MGVSVDLSIGHSFHSKICRASQYKLRKSSHLKLNRYSDGLLRF
GFSQDSGHKSSSKSGSEDFYIKVQKARGSHHPDYTALSSDGLGGQGVDFDPKE
PPKLPFPNQLMEGSEKGAVERPAFKPVLPRSGAILHSSPESASHQLHPAPDPKEQ
ELKPLCSGALSDSGRNSMLPHTSTSSYQLOPLVTPGPTSRFGSAAHNTQGV
LQDSNMMSLKALSFSDGSKLGNKADKSPVSRPSTDECSIQLEOKLREGA
LQKQRSFEELASSLAYERPRCRDELSGPPKGNKLQASQKSQRAQVHLQ
VLQKQKRQRLQELSLKQDLLETKLRSYERKTSFGPALETQWEVCQKSGEIS
LQKQKESQTEEVNAKSEIILGLKAQKDTKGLGLELRTDLEGALRTKGLLEV
ENELKKNKEALREKYNLRLRAELRERQHQDMSSGFQERHLVWKEEKVQ
YKQLOQSYVMYQNRQLEKALQQLARGDSAGEPLEVDLEGADIPYEDIATBI"
BASE COUNT 380 a 522 c 554 g 236 t
ORIGIN

Query Match 58.5%; Score 374; DB 9; Length 1692;
Best Local Similarity 83.0%; Pred. No. 7.5e-42;
Matches 483; Conservative 0; Mismatches 99; Gaps 1;

QY 1 GAGGTGACGCCNAGGCTAGCAGATCTGGTCTCAAGGCACAGCTGAAGGACACGGG 60
|||||
Db 1210 GAGGTGACGCCNAGGCTAGCAGATCTGGTCTCAAGGCACAGCTGAAGGACACGGG 1269
|||||

QY 61 GCGAAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120
|||||
Db 1270 GCGAAGCTGGAGGGCTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 1329
|||||

QY 121 GGCCTGGAGCTGGAGGCTGTGTGAGATGAGCTGCAGCGCGCAAGAACGAGCGGAGCTG 180
|||||

```

Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,  
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vaquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstock,G. and Gibbs,R.

Direct Submission  
 2 (bases 1 to 85530)  
 Unpublished  
 Worley,K.C.

Direct Submission  
 Submitted (02-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 85530)  
 Worley,K.C.

Direct Submission  
 Submitted (18-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GVRV  
 Center clone name: CH230-72A23  
 ----- Summary Statistics  
 Sequencing vector: plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 3389 bases at least Q40  
 Consensus quality: 35475 bases at least Q30  
 Consensus quality: 36705 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 40 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1040: contig of 1040 bp in length  
 1041 1140: gap of unknown length  
 1141 2462: contig of 1322 bp in length  
 2463 2563: gap of unknown length  
 2563 3575: contig of 1013 bp in length  
 3575 3676: gap of unknown length  
 3676 4689: contig of 1014 bp in length  
 4689 4790: gap of unknown length  
 4790 6224: contig of 1434 bp in length  
 6224 6323: gap of unknown length  
 6323 7473: contig of 1149 bp in length  
 7473 7573: gap of unknown length  
 7573 9106: contig of 1533 bp in length  
 9106 9205: gap of unknown length  
 9205 10878: contig of 1673 bp in length  
 10878 10979: gap of unknown length  
 10979 12427: contig of 1449 bp in length  
 12427 12528: gap of unknown length  
 12528 13987: contig of 1460 bp in length  
 13987 14087: gap of unknown length

14088 15164: contig of 1077 bp in length  
 15165 15264: gap of unknown length  
 15265 16281: contig of 1017 bp in length  
 16282 16381: gap of unknown length  
 16382 18060: contig of 1679 bp in length  
 18061 20139: gap of unknown length  
 20139 20233: contig of 1979 bp in length  
 20233 21324: gap of unknown length  
 21324 21424: contig of 1085 bp in length  
 21425 23233: contig of 1809 bp in length  
 23234 25051: gap of unknown length  
 25051 25151: contig of 1718 bp in length  
 25152 26674: gap of unknown length  
 26675 28334: contig of 1523 bp in length  
 28335 28434: gap of unknown length  
 28435 29845: contig of 1411 bp in length  
 29846 31160: gap of unknown length  
 31161 31260: contig of 1215 bp in length  
 31261 32414: gap of unknown length  
 32415 34151: contig of 1737 bp in length  
 34152 35731: gap of unknown length  
 35732 35831: contig of 1480 bp in length  
 35832 37026: gap of unknown length  
 37027 38890: contig of 1195 bp in length  
 38891 38990: contig of 1764 bp in length  
 38991 40845: gap of unknown length  
 40846 40945: contig of 1855 bp in length  
 40946 42520: gap of unknown length  
 42521 44652: contig of 1575 bp in length  
 44653 44752: gap of unknown length  
 44753 47079: contig of 2032 bp in length  
 47080 47179: gap of unknown length  
 47180 49524: contig of 2345 bp in length  
 49525 52769: gap of unknown length  
 52770 52869: contig of 3145 bp in length  
 52870 55421: gap of unknown length  
 55422 55521: contig of 2552 bp in length  
 55522 58766: gap of unknown length  
 58767 61389: contig of 3245 bp in length  
 61390 61489: gap of unknown length  
 61490 65525: contig of 2523 bp in length  
 65526 70064: gap of unknown length  
 70065 70164: contig of 4036 bp in length  
 70165 73090: contig of 4439 bp in length  
 73091 73190: gap of unknown length  
 73191 77062: contig of 2926 bp in length  
 77063 77162: gap of unknown length  
 77163 85530: contig of 3872 bp in length.  
 Location/Qualifiers  
 1..85530  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-72A23"  
 BASE COUNT 21327 a 17928 c 18750 g 22171 t 5354 others  
 ORIGIN  
 Query Match 36.4%; Score 232.4; DB 2; Length 85530;  
 Best Local Similarity 65.0%; Pred. No. 5.2e-23;  
 Matches 316; Conservative 0; Mismatches 166; Indels 4; Gaps 4;  
 QY 148 GAGCTGACGACAGACAGACGAGCGAGGTGTGTCGGGAGAGGTAACCTGCTGGAG 207





TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein, and its expression is altered in multiple human tumors  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)  
MEDLINE 99199287  
PUBMED 10097140  
REFERENCE 2 (bases 1 to 1515)  
AUTHORS Ishii, H., Baffa, R., Numata, S.I., Murakumo, Y., Rattan, S., Inoue, H., Mori, M., Fidanza, V., Alder, H. and Croce, C.M.  
TITLE Direct Submission  
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer Institute, 2338 10th street, Philadelphia, PA 19107, USA  
FEATURES  
Location/Qualifiers  
1..1515  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
/clone="D14"  
1..1515  
/gene="FEZ1"  
1..1515  
/gene="FEZ1"  
/feature="FEZ1"  
/feature="alternatively spliced"  
/codon\_start=1  
/product="FEZ1"  
/protein\_id="AAD23837.1"  
/db\_xref="GI:4572470"  
/translation="MGSVSLISGHFSHKRASYKLRKSHLKLNRVSDGLLRFP  
GFSQDSHGKSSKMGKEDFFYIKVQARGSHHPDYALSSDGLGGQAGVDFDPT  
PKLMPFNQLEMSKEGAIVPTAFKPLPRSGAILHSSPESASHQHPADPKPEK  
ELKPLCSGALSDSRNSMSLPHSTSSVOLDPLVTPVGTFRFGSARNITQGV  
LQDSNMWIKALSFGKSHGNSKADKSPVRSPISTDECSIQLEOKLLEREGA  
LQKQRFEEKELASLAYERPRRCDELGEPGPKNGKQKQASQRAQVILHLQ  
VLQQRQLRQELSEMSQDLLETKLRSYREKTSFGPALEETOWEVCOKSEIS  
LLKQQLKQTEVNAKASEILKQLKTRKLEGLRTQDLGALRTGLGALVEVC  
ENELQSVVAMVQNRORLEKALQQLARGSDSEPLEVDLEGADIPYEDIIATEI"  
BASE COUNT 336 a 483 c 482 g 214 t  
ORIGIN  
Query Match 27.4%; Score 175.4; DB 9; Length 1515;  
Best Local Similarity 59.9%; Pred. No. 6.5e-15;  
Matches 312; Conservative 0; Mismatches 206; Indels 3; Gaps 1;  
QY 65 ACCTGGAGGCGCTGAGCTGAGGACCCAGGACCTGAGGGCGCCCTGCGCACCAGGGCC 124  
DB 995 ACCTGCGAGTACTGAGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1054  
QY 125 TGGAGCTGGAGCTCTGTGAGATGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 184  
DB 1055 TCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1114  
QY 185 GGGAGAGGTGAACCTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 244  
DB 1115 GCTTCGGCCCGCGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1174  
QY 245 CCGGCGACATGGGCGCGCCCA---CCTTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 301  
DB 1175 CCCTCTTGAAGCAGGAGCTGGAAGGAGTCCAGACGAGGAGTGAACGCCAAGGCTAGCGAGA 1234  
QY 302 TGGAGCGGCTGCGGCGCGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361  
DB 1235 TCTTGGTCTCAAGGACAGCTGGAAGGACACGCGGGGCAAGCTGGAGGCGCTGGAGCTGA 1294  
QY 362 GTTTCAGCATGAGCGGCTGTGTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421  
DB 1295 GGACCCAGGACCTGGAGGCGCGCTCGGACCAAGGCGCTGGAGCTGGAGGCTGTGAGA 1354  
QY 422 ACAGCTGAGCAGAGCTACGTGGCCATCTACAGCGGAGGAGGAGGAGGAGGAGGAGG 481  
DB 1355 ATGAGCTGAGCAGAGCTACGTGGCCATCTACAGCGGAGGAGGAGGAGGAGGAGG 1414  
QY 482 TGCAGCAGCTGGCAGCTGGGAGCAGCGCGGGGAGCCCTTGGAGGTTGACCTGGAGGGG 541

Db 1415 TGCAGCAGCTGGCAGCTGGGGACAGCGCGGGGAGCGCTTGGAGGTTGACCTGGAAGGG 1474  
QY 542 CTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 582  
DB 1475 CTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1515  
RESULT 12  
BC005855  
LOCUS Homo sapiens, clone IMAGE:2988487, mRNA, partial cds.  
DEFINITION BC005855  
ACCESSION BC005855.1 GI:13543390  
VERSION  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1612)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov  
COMMENT Contact: MGC help desk  
Email: cgaps-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Institute for Systems Biology  
http://www.systemsbio.org  
contact: amadan@systemsbiology.org  
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia Greene, Mark Ketteman and Anuradha Madan  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Plate: 3 Row: n Column: 17  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.  
FEATURES  
Location/Qualifiers  
1..1612  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2988487"  
/tissue type="Colon, adenocarcinoma"  
/clone\_lib="NIH MGC\_15"  
/lab\_host="DH10B-R"  
/notes="vector: pOTB7"  
<1..934  
/codon\_start=2  
/product="Unknown (protein for IMAGE:2988487)"  
/protein\_id="AAH05855.1"  
/db\_xref="GI:13543391"  
/translation="ARGHWQREREALREDCAAQRAQQLQLQVQLQKQKQ  
LQDPALQREQLRRCATFLERLEETKWEVCOKGEISLLKQKQKESQ  
AELVQKSELVALRVALRATLRVSEGRARGLQEARARELEACQSLQRHQS  
AEQUREKAGQLDAEAGLRPPVPATADPFLAEDEAKVORAAAGVGGSLRAQVER  
LRVLEORREREEQRDSFEGERLAWQAEQVIRYQKQLQHNLYIQMYRRNRQLEQEL  
QQLSLEARELADLGLAEOAPCICLEITATEI"  
BASE COUNT 348 a 458 c 539 g 267 t  
ORIGIN  
Query Match 23.2%; Score 148.2; DB 9; Length 1612;  
Best Local Similarity 54.9%; Pred. No. 3.1e-11;  
Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;  
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTCTGGTCTCAAGGACACAGCTGAAGGACGCGG 60  
DB 311 GAGCTGGTGCAGAGGGGCGAGCTGGTGGTCTCTGCGGGTGGCGCTGCGGAGGCCCT 370

```
QY 61 GCACAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 120
Db |||||
QY 371 GCTACGCTGCGGCTCAGTGAGGCGCTGCGGCGGCTCTACAGGAGCGCCCGGAGCTCGG 430
Db |||||
QY 121 GGCCTGGAGCTGGAGGCTCTGAGAAATGAGCTGACGGCGCAAGAAAGAGGCGGAGCTG 180
Db |||||
QY 431 GAGCTGGAGCTGGAAAGCGCTGTTCCAGGAGCTGACAGCGACACCGCGACGAAAGCTGAGCAG 490
Db |||||
QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGGAGCTGACAGGAGCTGCGGCGC----- 230
Db |||||
QY 491 CTGCGGAGAGAGCTGGGAGTGGATGCTGAGGCGCGCGGACTCCGGAGGCCCTCTGTG 550
Db |||||
QY 231 -----CCAGGCGCGCTTGGCGCGGACATGGGCGGCC 264
Db |||||
QY 551 CCACCTGCCACCGCTGACCACTTCTCTGCGCAGAGAGTGATGAGGCGCAAAAGTGACGCGG 610
Db |||||
QY 265 ACCTTCCCGGAGGAGCTGCTGCTGCTGAGCGGAGCTGAGCGGCTGCGGCGCGAGCTG 324
Db |||||
QY 611 GCAGCAGCGCGGCTTGGGCGGAGCTTGGCGGCGCAGGTGGAGCGATTGCGGCTGGAGCTG 670
Db |||||
QY 325 CGGAGGAGCGGCAAGGCGCATGACAGATGCTCTCGGCTTCCAGCATGAGCGGCTGCTG 384
Db |||||
QY 671 CAGCGGAGCGCGCGGCTGAGAGCAGCGGACAGCTTTGAGGGGAGCGGCTGCGC 730
Db |||||
QY 385 TGGAGGAGGAGAGGAGAGGTGATTGATACCAAGAAACAGCTGCAGCAGAGCTACGCTG 444
Db |||||
QY 731 TGGCAGGCGAGAGAGGAGCAGGTGATCGCTACCAAGAGCAGCTGCAGCACAACACTACATC 790
Db |||||
QY 445 GCATGTACAGCGGAAACAGCGCTGAGAGAGCGCTTGGAGAGCGCTGCGAGCTGGCGGAC 504
Db |||||
QY 791 CAGATGTATCCGCGCAACCGCAGCTAGACAGGAGCTGCAGCAGCTCAGCCTGAGCTG 850
Db |||||
QY 505 AGCGCGGCGGAGCGCTTGGAGGTTGACCTGG-----AAGGGCTGACATCCCTACGAG 558
Db |||||
QY 851 GAGGCGCGGAGCTCGTGACCTGGGCTTGGCGGAGCAGCGCCCTGCTGCTGCTGCTGAG 910
Db |||||
QY 559 GACATCATAGCCACTGAGATCTGAGGCGCTGCTGGAGAGGCGAGCTCTGGGACCTGGCA 618
Db |||||
QY 911 GAGATCACTGCTACTGAGATCTAGGCGCTTACGCGCTCAGCAACAGCTCTGTAGGAGCTCTGCA 970
Db |||||
QY 619 CTGGG 623
Db |||||
QY 971 GAGGG 975
```

```
RESULT 13
HSM805394
LOCUS HSM805394 2275 bp mRNA linear PRI 12-JUL-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp761K1711 (from clone DKFZp761K1711).
ACCESSION AL834338
VERSION AL834338.1 GI:21739969
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2275)
AUTHORS Anorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp761K1711) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..2275
source
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="DKFZp761K1711"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2275
/gene="DKFZp761K1711"
<1..1597
/gene="DKFZp761K1711"
/notes="similarity to FEZ1 (Homo sapiens), different
splicing pattern"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD39005.1"
/db_xref="GI:21739970"
/translation="EKILIRPTAFKPVLPKPRGAPSLPSFMGRPATLGSQSGSLTQL
FGPSSSSSSSSSSAADKPLAFSGWAGCPSPGTLSDSRNLSLSLPTYSTGAEPTT
SSPGHLPSHGSGRGLPGPVPDDEALLHCVLEKLRDREAELQRLDLSDEATMCO
RASPDSSSCGERSPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
AYEPRQHWQREAREALREDCAQAQRAQQLLOLVFQLOQKRLQDPAOLLOE
REQRRCATLREQRLEBGRLEBKWEVCNGSEISLLKQQLKESQELHVRQAEQKAGOL
ALRVLRARATLRVSEGRARGLQEAARELELEACQELQRHREAEQKAGOL
DAEAAGREPPVPATADPFLAASDEAKVAAAGVGSRLRAVERLRVLRERERR
GEORDSFEGERLAWQAEQVIRYQKQLQHNITQMYRRNRQLBEQLQQLSLEARE
LADLGLAEQAPCILEITATEI"
2236..2241
/polyA_signal
/gene="DKFZp761K1711"
2257
/polyA_site
/gene="DKFZp761K1711"
BASE COUNT 455 a 697 c 739 g 384 t
ORIGIN
Query Match 23.2%; Score 148.2; DB 9; Length 2275;
Best Local Similarity 54.9%; Pred. No. 2.9e-11;
Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTCAAGCACACGCG 60
Db |||||
QY 974 GAGCTGTGCGAAGGCGCAGCTGTGCTCTGCGGTGCGGCGGCTCGGAGGCGCCGT 1033
Db |||||
QY 61 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAG 120
Db |||||
QY 1034 GCTACGCTGCGGCTCAGTGAGGCGCTGCGGCGGCTCTACAGGAGCGCGCGAGCTCG 1093
Db |||||
QY 121 GGCCTGAGCTGGAGGCTCTGTGAGAAATGAGCTGAGCGCAAGAAAGAGCGGAGCTG 180
Db |||||
QY 1094 GAGCTGGAGCTGGAAAGCCTGTTCCAGGAGCTGAGCGCACACCGCGAGAGCTGAGCAG 1153
Db |||||
QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGCAGCTGACAGGAGCTGCGGCGC----- 230
Db |||||
QY 1154 CTGCGGAGAGAGCTGGGAGTGGATGCTGAGGCGCGGAGCTCCGGAGCGCCCTGTG 1213
Db |||||
QY 231 -----CCAGGCGCGCTTGGCGCGGACATGGGCGGCC 264
Db |||||
QY 1214 CCACCTGCCACCGCTGACCACTTCTCTGCGCAGAGTGTAGGCGCAAGTGAGCGG 1273
Db |||||
QY 265 ACCTTCCCGGAGGAGCTGCTGCTGCTGAGCGGAGCTGGAGCGGCTGCGGCGCGAGCTG 324
Db |||||
QY 1274 GCAGCAGCGCGGCTTGGGCGGAGCTTGGGCGCGGAGCTGCGGCGGAGCTGCGGCTGAGCTG 1333
Db |||||
QY 325 CGGAGGAGCGGCAAGGCGCATGACAGATGCTCTGGGCTTCCAGCATGAGCGGCTGCTG 384
Db |||||
QY 1334 CAGCGGAGCGCGCGGCGGAGTGGAGCAGCGGAGCTTTGAGGGGAGCGGCTGGCC 1393
Db |||||
QY 385 TGGAGGAGGAGAGAGGAGAGGTGATTGATACCAAGAAACAGCTGCAGCAGAGCTACGCTG 444
Db |||||
QY 1394 TGGCAGGAGAGAGGAGCAGGTGATCGCTACCAAGAGAGCTGCAGCACAACACTACATC 1453
Db |||||
QY 445 GCCATGTACAGCGGAAACAGCGCTGGAGAGGCGCTTGCAGCAGCTGGCGAGCTGGGAC 504
Db |||||
```

1454 CAGATGTACCGCGCAACCGGAGCTAGACAGGAGCTGCAGAGCTCAGCTCAGCTGAGCTG 1513  
505 AGCGCGCGGAGCCCTTGGAGTTGACCTGG-----AAGGGCTGACATCCCTACAG 558  
1514 GAGCGCGGAGCTCGCTGACCTGGCGCTGCGCGAGCGGCCCTGCTGATCTGCTGAG 1573  
559 GACATCATGACCTGAGATCTGAGGGCTGCTGGGAGGCGAGCTGCGGAGCTGCGCA 618  
1574 GAGATCATGCTACTGAGATCTAGGCGCTCAGCAACAGCTCTGTAGGAGCTCTGCA 1633  
619 CTGGG 623  
1634 GAGGG 1638

RESULT 14  
BC006212 2766 bp mRNA linear PRI 12-JUL-2001  
LOCUS Homo sapiens, clone MGC:2586 IMAGE:3161855, mRNA, complete cds.  
DEFINITION BC006212  
ACCESSION BC006212  
VERSION BC006212.1 GI:13623228  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Strasbourg, R.  
Submitted (09-APR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

NIH-MGC Project URL: <http://mgs.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgs@nhgri.nih.gov](mailto:nisc_mgs@nhgri.nih.gov)

Shvchenko, Y., Wecherby, A.D., Beckstrom-Sternberg, S.M.,

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,

Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,

Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,

Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,

Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL plate: 5 Row: k Column: 1

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, Genomescan gene prediction.

Location/Qualifiers

1. .2766

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="MGC:2586 IMAGE:3161855"

/tissue\_type="Brain, neuroblastoma"

/clone\_lib="NIH MGC 19"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

1138. .2088

/codon\_start=1

/product="Unknown (protein for MGC:2586)"

/protein\_id="AAH06212.1"

/db\_xref="GI:13623229"

/translation="MCQAEERQHWREAREALREDCAAQRAQRAQLQLQVQFL

QOKROLODDPAOLLOEROLERRCATLEORELGPRLBETKWEVCKGSEISLLKQ  
OLKESAEVLVOKGSELVALREARATLRVSEGRARGLOEARAELELEACSOEL  
QROERAEQRLERAGQLDABAGLRPEPPVPATAPFLLAESDEAKVQRAAGVGGSL  
RAVRERLURVLRERRRBEQORDSFEGERLAWAEQEVIRYQKLOHNTIQMRRNR  
OLQBELOQLSLELEARELADLGAEPACICLEITATEI"

BASE COUNT 545 a 873 c 883 g 465 t

ORIGIN

Query Match 23.2%; Score 148.2; DB 9; Length 2766;

Best Local Similarity 54.9%; Pred. No. 2.8e-11;

Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;

QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTCTGGTCTCAAGGCACAGCTGAAGGACACCGGG 60

Db 1465 GAGCTGTGAGAGGCGGAGCTGTGGTCTCGGGTGGCGCTCGGGAGGCCCGT 1524

QY 61 GGCAAGCTGGAGGCGCTGGAGCTGAGACCCAGACCTGGAGGCGGCTCGGCACCAAG 120

Db 1525 GCTACGCTGGGGTCACTGAGGGCGCTGCGGGGTCTACAGGAGGCGGCGGAGCTCGG 1584

QY 121 GGCTTGGAGCTGGAGCTGTGAGAAATGAGTGCAGCGCAAGAAAGACGAGCGAGCTG 180

Db 1585 GAGCTGGAGCTGGAAGCTGTTCACAGAGCTGCAGCGACACCCGCCAGGAAGCTGAGCAG 1644

QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGAGCTGCAGAGCTGCGGCG----- 230

Db 1645 CTGCGGAGAGAGCTGGGCGAGTTGGATGCTGAGCGGCGGAGCTCCGGGAGCCCTGTG 1704

QY 231 -----CCAGCCCGCTTGGCGCGCGACATCGGCGCGGCC 264

Db 1705 CCACCTGCCACCGCTGACCCCATCTCTCTGCGAGAGTGTATGAGGCCAAAGTGCAGCGG 1764

QY 265 ACCTTCCCGGAGGAGCTCCCTGCGCTGAGCGGAGCTGGAGCGGCTGCGGGCGGAGCTG 324

Db 1765 GCAGCAGCGGGGTGGGGCGAGCTTGGGGCGGAGGTTGCGGGTGGAGCTG 1824

QY 325 CGGAGGAGGCGGCAAGCGCATGACAGATGCTTCGGGCTTCAGAGCTGAGCGGCTCGTG 384

Db 1825 CAGCGGAGGCGCGCGGGGTGAGGAGCAGCGGAGCAGCTTTGAGGGGAGCGGCTGCGC 1884

QY 385 TGGAGGAGGAGAGAGGAGAGTGAATTCAGTACAGAAACAGCTGCAGCAGAGCTACGTG 444

Db 1885 TGGCAGGCGAGAGAGGAGCAGTGTATCCGTACCAAGACAGCTGCAGCAACATACATC 1944

QY 445 GCGATGTACAGCGGAGACCGGCTGGAGAGCGCTTCAGAGCTGAGCGCTGCGGAGCGGAC 504

Db 1945 CAGATGTACCGCGGCAACCGGCGAGCTAGAGCAGAGCTGCAGCAGCTCAGCTGGAGCTG 2004

QY 505 AGCGCGGGGAGCGCTTGGAGGTTGACCTGG-----AAGGGGCTGACATCCCTTACGAG 558

Db 2005 GAGCGCGGGAGCTGCTGACCTGGGCTGCGGAGCGGCGGCTGCTGCTGCTGAG 2064

QY 559 GACATCATAGCCACTGAGATCTGAGGGGCTCCCTGGGAAGCGGAGTCTGGGGAGCTGCA 618

Db 2065 GAGATCACTGCTACTGAGATCTAGGCGCTCAGCAACAGCTCTGTAGGAGCTCTGCA 2124

QY 619 CTGGG 623

Db 2125 GAGGG 2129

RESULT 15

AB058716

LOCUS

DEFINITION

AB058716

ACCESSION

AB058716

VERSION

AB058716.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

1 (sites)

AB058716 5733 bp mRNA linear PRI 05-JUN-2001

Homo sapiens mRNA for KIAA1813 protein, partial cds.

AB058716

AB058716.1 GI:14017842

Homo sapiens brain cDNA to mRNA, clone:ph00819b.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

AUTHORS
TITLE

Nagase, T., Nakayama, M., Nakajima, D., Kikuno, R. and Ohara, O.  
Prediction of the coding sequences of unidentified human genes. XX.  
The complete sequences of 100 new cDNA clones from brain which code  
for large proteins in vitro  
DNA Res. 8 (2), 85-95 (2001)  
21245130  
2 (bases 1 to 5733)  
Ohara, O., Nagase, T. and Kikuno, R.  
Direct Submission  
Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan [E-mail:cdna.info@kazusa.or.jp].  
URL: <http://www.kazusa.or.jp/huge>, Tel:81-438-52-3913,  
Fax:81-438-52-3914

**JOURNAL**

21245130  
2 (bases 1 to 5733)

Ohara, O., Nagase, T. and Kikuno, R.

Direct Submission  
Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute,  
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba  
292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp.  
URL: <http://www.kazusa.or.jp/huge>, Tel: 81-438-52-3913,  
Fax: 81-438-52-3914)

## FEATURES

```

location|qualifiers
1. .5733
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="ph00819b"
/tissue_type="brain"
/notes="vector:pluascriptII SK plus"

gene
1. .5733
/gene="K1AA1813"
<3052. .5073
CDS
/gene="K1AA1813"
/notes="Start codon is not identified"
/codon_start=1
/product="K1AA1813 protein"
/protein_id="BAB47442.1"
/db xref="GI:14017843"

```

gene

/gene="KIAA1813"

CDS  
<3052. .5073

/gene="KIAA1813"

```

/note="Start codon is not identified."

```

/codon\_start=1

```
/product="KIAA1813 protein"
```

```
/protein_id="BAB47442.1"
```

```
/db xref="GI:14017843"
```

/translation="ASATMATIOTLPVLEPAPAAATPOAPVWGS/VSSILISGRCPCE  
GPAPRRHHGPGPTFFHQDGLLGGVBAEPLCPAPVPPKAPVPTSFTYINDEFRTF  
SPSPSSDADARORAHNAHRLKPIPLVSGLEKKNMEKILIRTPAATVLPKPR  
GAPSLPGRSITLSSQSGSLTQLFQGPASSSSSSAAADKLPAFGSAGSGPCS  
GTLSDGMRNSLPTTYSQGAETPSGCGHJPSHGSGRGALPGARPGVTPGSHSD  
SGRSSSSGSLGRVAGGLGSGTSPDSSGCBRSPPPPPPSSDALLACHVLE  
GKLDRRAELQOLSLDNEATQCAVESRQRHWOREALREDCAAOARAQRAQQ  
LIQVQFQOEKQQLQDDFAQLQREQLRRCATLERQRELPRLESTKWEVCQK  
SGEISLILKQSQEALVQKGSSELVALREARLTVSEGRGALQUGOAAAREL  
ELEACSLQORHQSEALQREKQDAAAGLREPEVPPATADPFLLAESDAKRVQ  
AAAGVGSRLAQVRLVRERRRGEQDPSFGEERLWAKAEKQIVRYQKQLOLHN  
YIOMYRNRLOEOLSLLEAREALRDLGLABOAPCI CLEIETATEI"

## BASE COUNT

ORIGIN

Query Match 23.2%; Score 148.2; DB 9; Length 5733;

Best Local Similarity 54.9%; Pred. No. 2.4e-11;

Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;

QY	1	GAGGTGAACCCCAAGGCTACGGAGATCCTCGGTCTCAAGGCACAGCTGAAGGACACGCGG	60
Db	4450	GAGCTGTGCAGAAAGGCGACGAGCTGGTGGCTCTGCGGTGGCGCTGCGGAGGCGCGT	4509
QY	61	GGCAAGCTGAGGGCCCTGGAGCTGAGACCCAGGACTGAGGGCGCCCTGCGCACCAAG	120
Db	4510	GCTACGCTCGGGTCAGTGAGGGCCGTGCGCGGGTCTACAGGAGCGCCCGCAGTCTCG	4569
QY	121	GGCCTGAGCTTGAGGCTCTGTGAATGAGCTGCAGCGCAAGAAGAACGAGGCGGAGCTG	180
Db	4570	GAGCTGAGCTGGAAGCTTGTCCAGGAGCTGCAGCGCACCCGCCAGGAAGCTGAGCAG	4629
QY	181	CTCGGGAGAGGTGAACCTGTCTGGAGCAGGAGCTGCAGGAGCTGCGGGC-----	230
Db	4630	CTCGGGAGAAAGCTGGGCAGTTGGATGTGAGGCGCCGGACTCGGGAGCCCCCTGTG	4689
QY	231	-----CCAGGCCCCCTGGCCCGCGACATGGGGCCGCCCC	264
Db	4690	CCACCTGCCACCGTGACCCATTCTCTCGCGGAGAGTGATGAGGCCAAAGTGACGGG	4749
QY	265	ACCTTCCCGAGACGTCTCTGCTGCACGCGGAGCTTGAGCGGGTTCGGGGCCGAGCTG	324
Db	4750	GCACAGCCCGGGTGTGGGGCAGCTTTCGGGGCCAGGTGAGAGCAATTCGGGTGGAGCTG	4809

Qy	325	CGGAGGAGCGGCAAGGCCATGACACAGATGTCTCTCGGGCTTCAGCATGAGCGGCTCGTG	384
Db	4810	CAGCGGAGCGCGCGGGGTGAGGAGCAGCGGACAGCTTTTGAGGGGAGCGCTGGCC	4869
Qy	385	TGGAAGGAGGAGAAGGAGAGGTGATTCAGTACCAGAAAAGCTGCAGCAGAGCTACGTG	444
Db	4870	TGCGACGCGAGAGAAGGAGCAGGTGATCCGCTACCAGAAGCAGCTGCAGCAACACTACATC	4929
Qy	445	GCCATGTACCAGCGGAAACAGCGCCTTGAGAGAGGCCCTGCAGCAGCTGGCAGCTGGGGAC	504
Db	4930	CAGATGTACCGGCGCAACCGGCAGCTAGAGCAGGAGCTGCAGCAGCTCAGCCTGGAGCTG	4989
Qy	505	AGGCCCGGGAGCGCCTTGAGGTTGACCTGG-----AAGGGCTGACATCCCCCTACGAG	558
Db	4990	GAGCGCCGGGAGTTCGCTGACCTGGCGCTTGGCGAGCAGGCCCCCTGCATCTCCCTGGAG	5049
Qy	559	GACATCATAGCCACTGAGATCTGAGGGGCTGCTCGGAAGGCGAGTCTGGGGACCTGGCA	618
Db	5050	GAGATCACTGCTACTGAGATCTAGGSCCTCAGCAACCCAGCTCTGTGAGGAGCTCTGCCA	5109
Qy	619	CTGGG	623
Db	5110	GAGGG	5114

Search completed: June 15, 2003, 02:18:41  
Job time : 1667.94 secs





PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
XX cancer -  
XX  
XX Example 2; Fig 5B; 255pp; English.  
XX  
XX The present sequence represents the cDNA sequence of the human FEZ1 gene.  
CC FEZ1 is a tumour suppressor gene, located at chromosome location 8p22.  
CC Decreased or no expression of FEZ1 is detected in a variety of cancer  
CC cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
XX Sequence 5492 BP; 1137 A; 1704 C; 1565 G; 1086 T; 0 other;  
SQ

Query Match 100.0%; Score 639; DB 21; Length 5492;  
Best Local Similarity 100.0%; Pred. No. 2.1e-99;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGGTGAACCCAGGCTAGCAGATCTGGTCTCAAGGCACAGCTGAAGACACGGG 60  
DB 1321 GAGGTGAACCCAGGCTAGCAGATCTGGTCTCAAGGCACAGCTGAAGACACGGG 1380

OY 61 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120  
DB 1381 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 1440

OY 121 GCGCTGGAGCTGGAGCTGTGAGATGAGCTGCAGCGCAAGAACGAGCGGAGCTG 180  
DB 1441 GCGCTGGAGCTGGAGCTGTGAGATGAGCTGCAGCGCAAGAACGAGCGGAGCTG 1500

OY 181 CTGCGGGAGAGTGAACCTGTGAGCAGAGCTGCAGGAGCTCGGGCCCGCCAGCCGCG 240  
DB 1501 CTGCGGGAGAGTGAACCTGTGAGCAGAGCTGCAGGAGCTCGGGCCCGCCAGCCGCG 1560

OY 241 CTGCGCCGCGACATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCGCGGGGAG 300  
DB 1561 CTGCGCCGCGACATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCGCGGGGAG 1620

OY 301 CTGGAGCGGCTCGGGCGGAGCTGCGGGAGAGCGGCAAGCCATGACAGATGCTCTCG 360  
DB 1621 CTGGAGCGGCTCGGGCGGAGCTGCGGGAGAGCGGCAAGCCATGACAGATGCTCTCG 1680

OY 361 GCGTTCCAGCATGACGGCTCTGTGGAGGAGGAGAGGAGAGGTTGATTCAGTACCAG 420  
DB 1681 GCGTTCCAGCATGACGGCTCTGTGGAGGAGGAGAGGAGAGGTTGATTCAGTACCAG 1740

OY 421 AAACAGCTGCAGCAGAGCTACGTGGCCATGTACACAGCGGAACACAGCGCTTGAGAGGCC 480  
DB 1741 AAACAGCTGCAGCAGAGCTACGTGGCCATGTACACAGCGGAACACAGCGCTTGAGAGGCC 1800

OY 481 CTGACAGCTGGCAGCTGGGAGACAGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540  
DB 1801 CTGACAGCTGGCAGCTGGGAGACAGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 1860

OY 541 GCTGACATCCCTACGAGACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGC 600  
DB 1861 GCTGACATCCCTACGAGACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGC 1920

OY 601 GAGTCTGGGAGCTGGCACTGGGAGCGAGGGCTCTCTCCCG 639  
DB 1921 GAGTCTGGGAGCTGGCACTGGGAGCGAGGGCTCTCTCCCG 1980

Db 1921 GAGTCTGGGAGCTGGCACTGGGAGCGAGGGCTCTCTCCCG 1959

RESULT 2  
AAA64507  
XX AAA64507 standard; DNA; 9048 BP.  
XX  
XX AAA64507;  
XX  
XX 02-JAN-2001 (first entry)  
XX  
XX Nucleotide sequence comprising the human FEZ1 gene.  
DE  
XX Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200050565-A2.  
XX  
XX 31-AUG-2000.  
XX  
XX 25-FEB-2000; 2000WO-US04950.  
XX  
XX 25-FEB-1999; 99US-0121537.  
XX  
XX (UVJE-) UNIV JEFFERSON THOMAS.  
XX  
XX Croce CM, Ishii H;  
PI  
XX WPI; 2000-558396/51.  
DR  
XX New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
XX Claim 2; Fig 5A; 255pp; English.  
XX  
XX The present sequence comprises the human FEZ1 gene. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased or no  
CC expression of FEZ1 is detected in a variety of cancer cells. Expression  
CC of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts  
CC with tubulin, with microtubules, and with protein EPI-gamma.  
CC Post-translational phosphorylation and dephosphorylation modulates the  
CC effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are  
CC useful for inducing cells to proliferate. Compounds which modulate FEZ1  
CC association with tubulin are useful for alleviating tubulin hyper- or  
CC hypo- polymerisation disorders, such as those associated with aberrant  
CC initiation of mitosis, modulation of the initiation and rate of cell  
CC proliferation and cell growth, modulation of cell shape, cell rigidity,  
CC cell motility, rate and stage of cellular DNA replication, intracellular  
CC distribution of organelles, metastatic potential of cell and cellular  
CC transformation from a non-cancerous to cancerous phenotype. Compounds  
CC which modulate FEZ1 binding and phosphorylation are also useful for  
CC alleviating a disorder, such as tumorigenesis, tumour survival, growth  
CC and metastasis.  
XX  
XX Sequence 9048 BP; 2011 A; 2605 C; 2490 G; 1939 T; 3 other;  
SQ

Query Match 100.0%; Score 639; DB 21; Length 9048;  
Best Local Similarity 100.0%; Pred. No. 2.1e-99;  
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGGTGAACCCAGGCTAGCAGATCTGGTCTCAAGGCACAGCTGAAGACACGGG 60  
DB 4912 GAGGTGAACCCAGGCTAGCAGATCTGGTCTCAAGGCACAGCTGAAGACACGGG 4971

OY 61 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 120  
DB 4972 GGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCCACCAAG 5031



QY 121 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAGAAAGAGCGAGCTG 180  
Db 5032 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAGAAAGAGCGAGCTG 5091  
QY 181 CTGCGGGAGAGGTGAACCTGCTGAGACAGAGCTGCAGAGAGCTCGGGGCCAGCGCCG 240  
Db 5092 CTGCGGGAGAGGTGAACCTGCTGAGACAGAGCTGCAGAGAGCTCGGGGCCAGCGCCG 5151  
QY 241 CTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGAG 300  
Db 5152 CTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGAG 5211  
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCG 360  
Db 5212 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCG 5271  
QY 361 GGCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGAGTGTATTCAGTACCAG 420  
Db 5272 GGCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGAGTGTATTCAGTACCAG 5331  
QY 421 AAACAGCTGCAGCAGCTACGTGGCCATGTACAGCGGAAACAGCGCTGGAGAGGCC 480  
Db 5332 AAACAGCTGCAGCAGCTACGTGGCCATGTACAGCGGAAACAGCGCTGGAGAGGCC 5391  
QY 481 CTGACGAGCTGGCAGCTGGGAGACAGCGCGGGAGCGCCCTTGGAGGTTGACCTGGAAGGG 540  
Db 5392 CTGACGAGCTGGCAGCTGGGAGACAGCGCGGGAGCGCCCTTGGAGGTTGACCTGGAAGGG 5451  
QY 541 GCTGATCCCTTACGAGACATATAGCACTGAGATCTGAGGGCTGCTGGGAAGGC 600  
Db 5452 GCTGATCCCTTACGAGACATATAGCACTGAGATCTGAGGGCTGCTGGGAAGGC 5511  
QY 601 GAGTCTGGGACCTGGCACTGGGAGCGAGGCTCTCCCG 639  
Db 5512 GAGTCTGGGACCTGGCACTGGGAGCGAGGCTCTCCCG 5550

RESULT 3

AAA64509  
ID AAA64509 standard; cDNA; 1791 BP.

XX AAA64509;

DT 02-JAN-2001 (first entry)

XX cDNA sequence encoding a human FEZ1 polypeptide.

XX Human; FEZ1 gene; tumour suppressor gene; 9p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT FH 1..1791  
FT CDS /\*tag= a  
FT /product= "FEZ1"

XX WO200050565-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

XX (UWJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Iehli H;

XX

DR WPI; 2000-558396/51.  
DR P-PSDB; AAB08715.  
XX  
PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
XX cancer.  
PS Claim 7; Fig 5I; 255pp; English.  
XX  
CC The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a  
CC tumour suppressor gene, located at chromosome location 9p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;  
Query Match 91.1%; Score 582; DB 21; Length 1791;  
Best Local Similarity 100.0%; Pred. No. 8.6e-90;  
Matches 582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60  
Db 1210 GAGGTGAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 1269  
QY 61 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 120  
Db 1270 GCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAG 1329  
QY 121 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAGAAAGAGCGGAGCTG 180  
Db 1330 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAGAAAGAGCGGAGCTG 1389  
QY 181 CTGCGGAGAGAGGTGAACCTGCTGAGCAGAGCTGCAGGAGCTCGGGGCCAGCGCCG 240  
Db 1390 CTGCGGAGAGAGGTGAACCTGCTGAGCAGAGCTGCAGGAGCTCGGGGCCAGCGCCG 1449  
QY 241 CTGCGCCCGCAGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCTGCAGCGGAG 300  
Db 1450 CTGCGCCCGCAGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCTGCAGCGGAG 1509  
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCG 360  
Db 1510 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGCTCTCG 1569  
QY 361 GGCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGAGAGAGAGTGTATTCAGTACCAG 420  
Db 1570 GGCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGAGAGAGAGTGTATTCAGTACCAG 1629  
QY 421 AAACAGCTGCAGCAGCTACGTGGCCATGTACAGCGGAAACAGCGCTGGAGAGGCC 480  
Db 1630 AAACAGCTGCAGCAGCTACGTGGCCATGTACAGCGGAAACAGCGCTGGAGAGGCC 1689  
QY 481 CTGAGCAGCTGGCAGCTGGGAGACAGCGCCGGGAGGCCCTTGGAGGTTGACCTGGAAGGG 540  
Db 1690 CTGAGCAGCTGGCAGCTGGGAGACAGCGCCGGGAGGCCCTTGGAGGTTGACCTGGAAGGG 1749  
QY 541 GCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 582  
Db 1750 GCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1791

```

RESULT 4
AAA64515
ID AAA64515 standard; cDNA; 1722 BP.
XX
AC AAA64515;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of truncated FEZ1 transcript G3612.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EF1-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 1..1722
FT /*tag= a
FT /product= "truncated FEZ1"
XX
FN WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Croce CM, Ishii H;
XX
DR WPI: 2000-558396/51.
DR P-PSDB; AAB08722.
XX
XX
PT New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
PS Disclosure; Fig 5H; 255pp; English.
XX
CC The present sequence encodes a truncated human FEZ1 polypeptide. The
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC suppressor gene, located at chromosome location 8p22. Decreased
CC expression of FEZ1 is detected in a variety of cancer cells.
CC or no expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC Expression of FEZ1 interacts tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC EF1-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;

Query Match 67.9%; Score 434; DB 21; Length 1722;
Best Local Similarity 88.1%; Pred. No. 8e-65;
Matches 513; Conservative 0; Mismatches 0; Indels 69; Gaps 1;

OY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCACACAGCTGAAGCACACGCGG 60
Db 1210 GAGGTGAACGCCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCACACAGCTGAAGCACACGCGG 1269
Oy 61 GGCNAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGGCTCGCGCACCAAG 120
Db 1270 GGCNAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGGCTCGCGCACCAAG 1329
Oy 121 GGCCTGGAGCTGGAGGCTCTGTGAGATGAGCTGAGCGCAAGAGAAACAGAGCGGAGCTG 180
Db 1330 GGCCTGGAGCTGGAGGCTCTGTGAGATGAGCTGAGCGCAAGAGAAACAGAGCGGAGCTG 1389
Oy 181 CTGCGGAGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCGCCAGCGGCC 240
Db 1390 CTGCGGAGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCGCCAGCGGCC 1449
Oy 241 CTGCGCGCGACATGGGGCGCGCCACCTTCCCGAGGACGCTCCCTGCGCTGCAGCGGAG 300
Db 1450 CTGCGCGCGACATGGGGCGCGCCACCTTCCCGAGGACGCTCCCTGCGCTGCAGCGGAG 1509
Oy 301 CTGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGCAAGGCCATGACCATGTCTCTCG 360
Db 1510 CTG-----
Oy 361 GGCCTCCAGCATGAGCGGCTCGTGTGAAGGAGGAGAGAGAGAGGTGATTTCAGTACCAG 420
Db 1513 -----GAGCGGCTCGTGTGAAGGAGGAGAGAGAGAGGTGATTTCAGTACCAG 1560
Oy 421 AAACAGCTGCAGCAGAGCTAGCTGGCCATGTACAGCGGAAACAGCGGCTCGGAGAGGCC 480
Db 1561 AAACAGCTGCAGCAGAGCTAGCTGGCCATGTACAGCGGAAACAGCGGCTCGGAGAGGCC 1620
Oy 481 CTGAGCAGCTGGCAGCTGGGGAGCAGCGCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 540
Db 1621 CTGAGCAGCTGGCAGCTGGGGAGCAGCGCGGGAGCGCTTGGAGGTTGACCTGGAAGGG 1680
Oy 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 582
Db 1681 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1722

RESULT 5
AAA64514
ID AAA64514 standard; cDNA; 1692 BP.
XX
AC AAA64514;
XX
DT 02-JAN-2001 (first entry)
XX
DE Nucleotide sequence of truncated FEZ1 transcript G3611.
XX
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW tumour proliferation; tubulin; microtubule; protein EF1-gamma;
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW tumorigenesis; tumour survival; metastasis; ss.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT 1..1692
FT /*tag= a
FT /product= "truncated FEZ1"
XX
FN WO200050565-A2.
XX
PD 31-AUG-2000.
XX
PF 25-FEB-2000; 2000WO-US04950.
XX
PR 25-FEB-1999; 99US-0121537.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
```

PI Croce CM, Ishii H;  
XX WPI; 2000-558396/51.  
DR P-PSDB; AAB08721.  
XX  
PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
PS  
PS Disclosure; Fig 5G; 255pp; English.  
XX  
XX The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EFl-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
XX Sequence 1692 BP; 380 A; 523 C; 553 G; 236 T; 0 other;  
SQ

Query Match 58.5%; Score 374; DB 21; Length 1692;  
Best Local Similarity 83.0%; Pred. No. 1.1e-54;  
Matches 483; Conservative 0; Mismatches 0; Indels 99; Gaps 1;

QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACACGCTGAAGGACACGCGG 60  
Db 1210 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACACGCTGAAGGACACGCGG 1269

QY 61 GGCAAGCTGGAGGGCTGGAGTGAAGACCTGAGGACCTGGAGGGCGCCCTGGCGACCAAG 120  
Db 1270 GGCAAGCTGGAGGGCTGGAGTGAAGACCTGAGGACCTGGAGGGCGCCCTGGCGACCAAG 1329

QY 121 GGCCTGGAGCTGGAGTCTGTGAGATGAGCTGACGCGCAAGAACGAGCGGAGCTG 180  
Db 1330 GGCCTGGAGCTGGAGTCTGTGAGATGAGCTGACGCGCAAGAACGAGCGGAGCTG 1389

QY 181 CTGCGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCCCGCCAGCGCC 240  
Db 1390 CTGCGGGAGAGGTGAACCT----- 1409

QY 241 CTGCGGGCGGACATGGGGCGGCCACCTTCCCGAGGAGCTCCCTCGCCCTGGCAGGGAG 300  
Db 1410 -----G 1410

QY 301 CTGGAGCGGCTCGGGCGGAGCTGCGGGAGGAGCGCAAGGCATGACACAGATGTCCTCG 360  
Db 1411 CTGGAGCGGCTCGGGCGGAGCTGCGGGAGGAGCGCAAGGCATGACACAGATGTCCTCG 1470

QY 361 GGCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGGAGGTGATTCAGTACCAG 420  
Db 1471 GGCCTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGGAGGAGGTGATTCAGTACCAG 1530

QY 421 AAACAGCTGCACAGAGTACTGTGGCCATGTACACGCGGAACACGCGCTGGAGAGGCC 480  
Db 1531 AAACAGCTGCACAGAGTACTGTGGCCATGTACACGCGGAACACGCGCTGGAGAGGCC 1590

QY 481 CTGCAGCAGCTGGCAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540  
Db 1591 CTGCAGCAGCTGGCAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 1650

OY 541 GCTGACATCCCTACGAGGACATCATGACCCACTGAGATCTGA 582  
Db 1651 GCTGACATCCCTACGAGGACATCATGACCCACTGAGATCTGA 1692

RESULT 6  
AAA64510  
ID AAA64510 standard; cDNA; 404 BP.  
XX AC AAA64510;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Nucleotide sequence of truncated FEZ1 transcript E16T0.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EFl-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..231  
FT /\*tag= a  
FT /product= "truncated FEZ1"  
XX  
PN WO200050565-A2.  
XX  
PD 31-AUG-2000.  
XX  
PF 25-FEB-2000; 2000WO-US04950.  
XX  
PR 25-FEB-1999; 99US-0121537.  
XX  
PA (UYJE-) UNIV JEFFERSON THOMAS.  
XX  
PI Croce CM, Ishii H;  
XX  
XX WPI; 2000-558396/51.  
DR P-PSDB; AAB08717.  
XX  
PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
XX Disclosure; Fig 5C; 255pp; English.  
XX  
CC The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EFl-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 404 BP; 92 A; 119 C; 127 G; 66 T; 0 other;  
Query Match 39.0%; Score 249; DB 21; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.3e-33;

Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 CGGCAAGCCATGACAGATGCTCTCGGGCTTCCAGCATGAGCGCTGTGTGGAGGAG 393  
 |||||  
 DB 156 CGGCAAGCCATGACAGATGCTCTCGGGCTTCCAGCATGAGCGCTGTGTGGAGGAG 215  
 |||||  
 QY 394 GAGAAGGAGAGGTGATTTCAGTACCAAGAAACAGCTGACAGAGTACGTGGCCATGTAC 453  
 |||||  
 DB 216 GAGAAGGAGAGGTGATTTCAGTACCAAGAAACAGCTGACAGAGTACGTGGCCATGTAC 275  
 |||||  
 QY 454 CAGCGGAACACGCGCTGGAGAGCCCTTGACAGAGCTGGCAGCTGGGGACAGCCCGGG 513  
 |||||  
 DB 276 CAGCGGAACACGCGCTGGAGAGCCCTTGACAGAGCTGGCAGCTGGGGACAGCCCGGG 335  
 |||||  
 QY 514 GAGCCCTGGAGGTGACCTGAAAGGGCTGACATCCCTACGAGGACATATAGCCACT 573  
 |||||  
 DB 336 GAGCCCTGGAGGTGACCTGAAAGGGCTGACATCCCTACGAGGACATATAGCCACT 395  
 |||||  
 QY 574 GAGATCTGA 582  
 |||||  
 DB 396 GAGATCTGA 404  
 |||||

RESULT 7

AAA64512  
 ID AAA64512 standard; cDNA; 1614 BP.  
 XX  
 AC AAA64512;

XX 02-JAN-2001 (first entry)

DE Nucleotide sequence of truncated FEZ1 transcript T8D145M4.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
 KW tumour proliferation; tubulin; microtubule; protein FE1-gamma;  
 KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
 KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
 KW tumorigenesis; tumour survival; metastasis; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 1..633  
 CDS /\*tag= a  
 FT /\*product= "truncated FEZ1"

XX WO200050565-A2.

XX 31-AUG-2000.

XX 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Ishii H;

XX WPI; 2000-558396/51.

XX P-PSDB; AAB08719.

XX New polynucleotide homologous with a portion of one strand of the human  
 PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
 PT cancer -

XX Disclosure; Fig 5B; 255pp; English.

XX The present sequence encodes a truncated human FEZ1 polypeptide. The  
 CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
 CC suppressor gene, located at chromosome location 8p22. Decreased  
 CC or no expression of FEZ1 is detected in a variety of cancer cells.  
 CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
 CC also interacts with tubulin, with microtubules, and with protein

CC EF1-gamma. Post-translational phosphorylation and dephosphorylation  
 CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
 CC expression are useful for inducing cells to proliferate. Compounds  
 CC which modulate FEZ1 association with tubulin are useful for alleviating  
 CC tubulin hyper- or hypo- polymerisation disorders, such as those  
 CC associated with aberrant initiation of mitosis, modulation of the  
 CC initiation and rate of cell proliferation and cell growth, modulation of  
 CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
 CC DNA replication, intracellular distribution of organelles, metastatic  
 CC potential of cell and cellular transformation from a non-cancerous to  
 CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
 CC phosphorylation are also useful for alleviating a disorder, such as  
 CC tumorigenesis, tumour survival, growth and metastasis.

XX Sequence 1614 BP; 367 A; 500 C; 522 G; 225 T; 0 other;

Query Match 35.1%; Score 224.4; DB 21; Length 1614;  
 Best Local Similarity 65.9%; Pred. No. 1.9e-29;  
 Matches 325; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 90 CCAGGACCTGAGGGCGCCCTGCGCACCAGGGCTGGAGCTGTGTGAGAATGA 149  
 |||||  
 DB 1122 CCGCGCGCTGGAGGAGACCCAGTGGAGGTGTCCAGAAAGTCAGCGGAGATCTCCCTCCT 1181  
 |||||  
 QY 150 GCTGCGAGCGCAAGAAACAGAGCGGAGCTGCTCGGGAGAGAGGTGAACCTGCTGGAGCA 209  
 |||||  
 DB 1182 GAAGCAGAGCTGAAGGAGTCCAGACGAGAGGTGAACGCCAAGGCTAGCGAGATCCTGGG 1241  
 |||||  
 QY 210 GGAGCTGAGAGCTGCGGGCCCGAGCGCCCTGCGCCCGCGACATGGGGCGGCCACCTT 269  
 |||||  
 DB 1242 TCTCAAGGCACAGCTGAAGGACACGCGGGGCAAGCTGGAGGCGCTGGAGCTGAGACCCA 1301  
 |||||  
 QY 270 CCGCGAGACCTCCCTGCGCTGCGAGCGGAGCTGGAGCGGCTGCGGGCGGAGCTCGGGA 329  
 |||||  
 DB 1302 GGAGCTGGAGGGCGCCCTGCGCACCAAGGGCTGGAGCTGGAGGTCTGTGAGAATGAGCT 1361  
 |||||  
 QY 330 GGAGCGGCAAGGCCATGACACAGATGCTCCTCGGGCTTCCAGCATGAGCGGCTCGTGTGAA 389  
 |||||  
 DB 1362 GCAGCGCAAGAAACAGAGCGGAGCTGCTCGGGAGAGAGCATGAGCGCTCGTGTGAA 1421  
 |||||  
 QY 390 GGAGGAGAGGAGAGGTGATTTCAGTACCAAGAAACAGCTGCGAGCAGAGCTAGTGGCCAT 449  
 |||||  
 DB 1422 GGAGGAGAGGAGAGGTGATTTCAGTACCAAGAAACAGCTGCGAGCAGAGCTAGTGGCCAT 1481  
 |||||  
 QY 450 GTACCGCGGAACACAGCGCCCTGGAGAGCCCTGCGAGCAGCTGGCAGCTGGGAGCAGCGC 509  
 |||||  
 DB 1482 GTACCGCGGAACACAGCGCCCTGGAGAGCCCTGCGAGCAGCTGGCAGCTGGGAGCAGCGC 1541  
 |||||  
 QY 510 CCGGGAGCCCTTGGAGGTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGC 569  
 |||||  
 DB 1542 CCGGGAGCCCTTGGAGGTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGC 1601  
 |||||  
 QY 570 CACTGAGATCTGA 582  
 |||||  
 DB 1602 CACTGAGATCTGA 1614  
 |||||

RESULT 8

AAA64513  
 ID AAA64513 standard; cDNA; 1512 BP.

XX AAA64513;

XX 02-JAN-2001 (first entry)

XX Nucleotide sequence of truncated FEZ1 transcript D14.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
 KW tumour proliferation; tubulin; microtubule; protein FE1-gamma;  
 KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
 KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
 KW tumorigenesis; tumour survival; metastasis; ss.

OS Homo sapiens.  
XX Key Location/Qualifiers  
FH 1..1512  
FT /\*tag= a  
FT /product= "truncated FEZ1"  
XX WO200050565-A2.  
XX 31-AUG-2000.  
XX 25-FEB-2000; 2000WO-US04950.  
XX 25-FEB-1999; 99US-0121537.  
XX (UYJE-) UNIV JEFFERSON THOMAS.  
XX Croce CM, Ishii H;  
XX WPI; 2000-558396/51.  
DR P-PSDB; AAB08720.  
XX  
XX New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
XX Disclosure; Fig 5F; 255pp; English.  
XX  
XX The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EF1-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 1512 BP; 335 A; 483 C; 481 G; 213 T; 0 other;  
Query Match 27.0%; Score 172.4; DB 21; Length 1512;  
Best Local Similarity 59.7%; Pred. No. 1.1e-20;  
Matches 309; Conservative 0; Mismatches 206; Indels 3; Gaps 1;  
QY 65 AGCTGAGGCGCTGAGTGAGGAGCCAGGACCTGGAGGGCGCCCTGCGCACCAAGGGCC 124  
DB 995 ACCTCGAGTACTGACGCTTCAGCAGGAGAGCGGAGCTCGGAGGCTCGAGAGCC 1054  
QY 125 TGGAGCTGGAGTCTGTGAGATGAGTGCAGCGGCAAGAGAACAGGCGGAGCTGTGC 184  
DB 1055 TCATGAAGGAGCAGGAGACCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCA 1114  
QY 185 GGGAGAGGTGAACCTGTGGAGCAGGAGCTCAGGAGCTGGCGGCCAGGCGCCCTGG 244  
DB 1115 GCTTCGGCCCGCTGGAGGAGACCCAGTGGGAGGTGTGCAGAAAGTCAGCGGAGATCT 1174  
QY 245 CCCGCGACATGGGGCCGCCCA---CTTCCCGAGGAGCTCCCTCCCTGCGCGGGAGC 301  
DB 1175 CCCTCTGAAGCAGCAGCTGAAGGAGTCCAGACGAGGTGAACCCCAAGGCTAGCGAGA 1234  
QY 302 TGGAGCGCTGGGGCCGAGCTGCGGAGGAGCGCAGGCGCATCAGCATGTCTCTGG 361  
DB 1235 TCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGGGGCAAGCTGGAGGGCCTGGAGCTGA 1294

QY 362 GCTTCCAGCATGAGCGGCTCGTGTGAAGGAGGAGAGGAGGATTCAGTACCAGA 421  
DB 1295 GGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAGGGCTGGAGGTCTGTGAGA 1354  
QY 422 AACAGCTGCGAGCAGAGCTACCTGGCCATGTATCCAGCGGAAACAGCGCTCGAGAAAGGCC 481  
DB 1355 ATGAGCTGCGAGCAGAGCTACCTGGCCATGTATCCAGCGGAAACAGCGCTCGAGAAAGGCC 1414  
QY 482 TGCAGCAGCTGGCAGCTGGGACAGCGCGGGGAGCCCTTGGAGTTTGCACCTTGAAGGGG 541  
DB 1415 TGCAGCAGCTGGCAGCTGGGACAGCGCGGGGAGCCCTTGGAGTTTGCACCTTGAAGGGG 1474  
QY 542 CTGACATCCCTTACGAGGACATCATAGCCACTGAGATC 579  
DB 1475 CTGACATCCCTTACGAGGACATCATAGCCACTGAGATC 1512  
RESULT 9  
AAF15625  
ID AAF15625 standard; cDNA; 1617 BP.  
XX  
XX AAF15625;  
AC  
XX  
DT 13-MAR-2001 (first entry)  
XX  
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:60.  
XX  
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
XX wound; infectious disease; ss.  
OS Homo sapiens.  
XX WO200055174-A1.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US05988.  
XX 12-MAR-1999; 99US-0124270.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587513/55.  
XX P-PSDB; AAB56422.  
XX  
XX Prostate cancer associated gene sequences, referred to as prostate  
PT cancer antigens, useful for treatment, prevention, and diagnosis of  
PT disorders such as prostate cancer -  
XX  
XX Claim 1; Page 657; 2338pp; English.  
XX  
XX AAF15566 to AAF16505 encode the human prostate cancer associated  
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
CC The prostate cancer antigens can have neuroprotective, cytostatic,  
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAB57303 represent sequences used in the exemplification of the present  
CC invention.

XX SQ Sequence 1617 BP; 355 A; 457 C; 532 G; 268 T; 5 other;

Query Match 23.3%; Score 149.2; DB 21; Length 1617;  
 Best Local Similarity 55.0%; Pred. No. 9.1e-17;  
 Matches 365; Conservative 0; Mismatches 258; Indels 41; Gaps 2;

QY 1 GAGGTGAACGCCAAGCTAGCGAGATCTGGTCTCAAGGCACACGCTGAAGACACGCGG 60  
 DB 279 GAGCTGGTGCAGAGGCGAGCTGGTGGCTCTGCGGGTGGCGCTCGGAGGCGCCGT 338  
 QY 61 GCGAAGCTGGAGGCGCTGAGCTGAGGACCCAGAGACCTGGAGGCGCCCTGCGACCAAG 120  
 DB 339 GCTACGCTCGGGTCACTGAGGCGCGTGGCGGGTCTACAGGAGCGCCCGAGCTCG 398  
 QY 121 GCGCTGGAGCTGAGGCTCTGAGATGAGCTGAGCGCAAGAAAGAGCGGAGCTG 180  
 DB 399 GAGCTGGAGCTGGAAGCTGTTCCAGGAGCTGACAGCGACACCGCCAGGAAGCTGAGCAG 458  
 QY 181 CTGCGGAGAGGTGAACCTGCTGAGCAGGAGCTGCAGGAGCTGAGGAGCTGCGGCGC----- 230  
 DB 459 CTGCGGAGAAAGCTGGGAGCTGGATGCTGAGGCGCGCGGAGCTCCGGAGCGCCCTCTGTG 518  
 QY 231 -----CCAGGCGCGCCCTGGCGCGGACATGGGCGCGCC 264  
 DB 519 CCACTGCCACCGCTGACCCATTCTCTGGCAGAGAGTGATGAGGCCAAGTGAGCGG 578  
 QY 265 ACCTTCCCGAGGAGCTCCCTGCTGAGCGGGAGCTGGAGCGCTCGGGCGGAGCTG 324  
 DB 579 GCAGCAGCGCGGTTGGGGGCGAGCTTGGCGGCGGAGCTGGAGCGATTCGGGTGAGCTG 638  
 QY 325 CGGAGGAGCGCAGGCGCATACAGATGCTCTCGGCTTCCAGCATGAGCGGCTCTGTG 384  
 DB 639 GAGCGGAGCGCGCGGGGTGAGGAGCAGCGGAGCAGCTTTGAGGGGAGCGGCTGGCC 698  
 QY 385 TGAAGAGGAGGAGGAGGAGGATTCAGTACCAAGAAACAGCTGCAGCAGAGCTACCTG 444  
 DB 699 TGGCAGGAGAGAGGAGGAGGAGTATCGCTACCAAGAGCAGCTGCAGCAGACTACATC 758  
 QY 445 GCATGTACCGAGGAGCAGCGCTGAGGAGGCGCTGAGCAGCTGCAGCTGGCGGAGC 504  
 DB 759 CAGATGTATACCGGCGAACCGGAGCTAGAGCAGGAGCTGCAGCAGCTCAGCTGAGCTG 818  
 QY 505 ACGCGCGGGAGCCCTGAGGCTGACCTG-----GAGGGGCTCACATCCCTACGAGG 559  
 DB 819 GAGGCGCGGAGCTGCTGACCTGGGCTGGCGGAGAGCCCTGCTATCTGCTGGAGG 878  
 QY 560 ACATCATAGCCACTGAGATCTGAGGGGCTGCGTGGGAGGCGAGTCTGGGGACCTGGCAC 619  
 DB 879 AGATCACTGCTACTGAGATCTAGGCGCCCTCAGCAACCCAGCTCTGTAGGAGCTCTGCCAG 938  
 QY 620 TGGG 623  
 DB 939 AGGG 942

RESULT 10  
 ID ABQ54361 standard; cDNA; 1623 BP.  
 XX AC ABQ54361;  
 XX AC  
 XX 22-AUG-2002 (first entry)  
 XX DE Human ovarian antigen HELGN26 cDNA, SEQ ID NO:241.  
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; gene; ss.  
 OS Homo sapiens.  
 PN WO200200677-A1.  
 XX 03-JAN-2002.  
 PD 07-JUN-2001; 2001WO-US18569.  
 PF 07-JUN-2000; 2000US-209467P.  
 PR (HUMA-) HUMAN GENOME SCI INC.  
 PA Birse CE, Rosen CA;  
 XX WPI; 2002-147878/19.  
 PI P-PSDB; ABP41284.  
 DR Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 XX useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and  
 PT neurological diseases -  
 XX Claim 1; SEQ ID No 241; 2922pp; English.  
 PS The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 XX ABP43228) and to cDNAs encoding them (ABQ54361-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovarian and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC and respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents cDNA encoding a human ovarian antigen of the  
 CC invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 1623 BP; 358 A; 457 C; 534 G; 268 T; 6 other;

Query Match 23.2%; Score 148.2; DB 24; Length 1623;  
 Best Local Similarity 54.9%; Pred. No. 1.3e-16;  
 Matches 365; Conservative 0; Mismatches 258; Indels 42; Gaps 2;

QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCTGGTCTCAAGGCACACGCTGAAGACACGCGG 60  
 DB 282 GAGCTGGTGCAGAGGCGGAGCTGGTGGCTCTGCGGGTGGCGCTCGGAGGCGCCGT 341  
 QY 61 GCGAAGCTGGAGGCGCTGAGCTGAGGACCCAGAGCTGGAGGCGCCCTCGGACCAAG 120  
 DB 342 GCTACGCTGGGGTCACTGAGGCGCGCTGCGGGGCTCTACAGGAGGCGCGCGGAGCTCGG 401

Qy	121	GGCCTGGAGCTGGAGGTCTGTGAGAAATGAGTGCAGCGCAAGAAAGACGAGCGGAGCTG	181
Db	402	GAGCTGGAGCTGGAAGCTGTGTCCAGGAGCTGCAGCGACACCGCAGGAAGCTGAGCAG	461
Qy	181	CTGGGGGAGAGGTGAACCTCTCTGAGCAGGAGCTGCAGGAGCTTGGGGC-----	230
Db	462	CTGGGGGAGAAAGCTGGGCAAGTTGGATGCTGAGCGCGCGGAGCTCCGGGAGACCCCTGTG	521
Qy	231	-----CCAGGCGCGCCTGGCGCCGGACATGCGGCGCGCATGCGGCGCGCC	264
Db	522	CCACCTGCCACCGCTGACCCATTCTCTCTGGCAGAGAGTGATGAGGCCAAAGTGACGCG	581
Qy	265	ACCTTCCCGAGAGAGCTCCCTGCTCTGAGCGGGAGCTGGAGCGGCTGCGGGCCGAGCTG	324
Db	582	GCAGCAGCCGGGTGTGGGGGAGAGCTTGGCGGCCACAGTGGAGCGATTGCGGTGGAGCTG	641
Qy	325	CGGGAGGAGCGCAAGGCCATGACACAGATGTCTCGGGCTTCCAGCATGAGCGGCTCGTG	384
Db	642	CAGCGGAGCGCGCGGCGTGAGGAGCAGCGGGACAGCTTTGAGGGGAGCGGCTGGCC	701
Qy	385	TGGAAGGAGGAGAGAGGAAGGTGATTCAGTACCAGAAACAGCTGCAGCAGAGCTACGCTG	444
Db	702	TGGCAGGCAGAGAGGACAGGTGATCCGCTACCAGAGCAGCTGCAGCACAACTACATC	761
Qy	445	GCCATGTACACGCGAACCAGCGCTGAGAGAGCCCTGCAGCAGCTGSCACGTGGGAGC	504
Db	762	CAGATGTATCCGGCGCAACCGGCAGCTAGACAGGAGCTGCAGCAGCTCAGCGCTGAGCTG	821
Qy	505	AGCCCCGGGAGCCCTTGGAGGTTGACCTGG-----AAGGGGCTGACATCCCTACGAG	558
Db	822	GAGGCCCGGGAGCTCGCTGACCTGGGCTTGGCCGAGCAGGCCCTCTGCATCTGCTGGAG	881
Qy	559	GACATCATAGCCACTGAGATCTGAGGGGCTGCTGGGAAGGGAGTCTTGGGACACTGGCA	618
Db	882	GAGATCACTGTACTGAGATCTTAGGGCCCTCAGCAACCAAGCTCTGTAGGGAGCTCTGCCA	941
Qy	619	CTGGG	623
Db	942	GAGGG	946

RESULT 11

AAA50254  
ID AAA50254 standard; DNA; 1926 BP.

AA  
AC AAA50254;

DT 07-NOV-2000 (first entry)

DE Epstein Barr virus nuclear antigen 1 protein (EBNA1) DNA.

EBV; nuclear antigen 1; EBNA1; episome; transfection; selection;  
KW gene therapy; ds.  
KW

OS Epstein-barr virus

XX PN WO2000047778-A1

XX 17-AUG-2000

XX  
DE 11 FEB 2000 0000Z 430354Z

XX

XX  
XXXX  
XX  
FIFTH COFFEE INC.XX  
FI HORLICK KA, Chetlsky D;

DR: WPI; 2000-515062/46.  
DR: P-DCDB: AAYQ595C

[illegible]

PT production of a desired protein in vitro and for gene therapy -

Disclosure; Fig 2; 53pp; English.

The present sequence is that of DNA encoding the Epstein-Barr virus (EBV) nuclear antigen protein 1 (EBNA1, see AA095956). EBNA1 is utilised in a novel method for obtaining a eukaryotic cell that is stably transfected with at least one episome. This method involves transfecting a eukaryotic cell with: (1) a first episome comprising an EBV origin of replication (oriP, see AA050253), a gene encoding a first protein whose expression results in cell death and a selectable marker for eukaryotic cells; and (2) a second episome comprising an EBV oriP and a gene encoding a second protein, where expression of the second protein prohibits the occurrence of cell death resulting from expression of the first protein to produce doubly transfected cells which also express an antigen that promotes retention of the episomes by the cells. The doubly transfected cells are maintained under conditions in which the first and second proteins and the selectable marker are expressed, and the selective pressure specified by the marker is maintained. Under these conditions, only cells containing both episomes live. Preferably, EBNA1 is expressed from 1 of the episomes, and the protein of interest from the other episome. Either or both epitopes may further comprise a nucleic acid sequence encoding a protein desired to be expressed in the cell (e.g. a therapeutic protein), a nucleic acid encoding an RNA that is not intended to be translated (e.g. a therapeutic RNA), or a DNA sequence used as a tag for the cells. The method is applicable to cell culture or intact organisms, for gene therapy. It allows the rapid establishment of eukaryotic cells that stably and reliably express a gene of interest, using a novel method of selection, and maintenance of that selection without the need for exogenous selection factors, such as antibiotics.

Sequence 1926 BP; 487 A; 352 C; 872 G; 215 T; 0 other;

Query Match 17.9%; Score 114.4; DB 21; Length 1926;

Query Match 17.3%; Score 114.4; DB 2  
Best Local Similarity 49.8%; Pred. No. 6.8e-11;

DDB: Local Similarity 45.0%, P-Val: NO. 0.00-11,  
 Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

Qy	1	GAGGTGAACCGCAGCGCTAGCGAGATCTCTGGTCTCAAGGCACAGCTTAAGGACACCGCG	60
Db	464	GAGGGCAGGACAGGAGGAGGGGAGGAGGGGAGGAGGGGAGGAGGAGGAGGGG	523
Qy	61	GGCAAGCTGAGGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTTCGCGACCAAG	120
Db	524	CAGGAGCAGGAGGGGCGAGGGGCGAGGACAGGAGGAGGGGCGAGGAGGGCGAGG	583
Qy	121	GGCTCTGAGCTGGAGGTCTGTGAGAACTGAGTCAGCGCAAGAAAGACGAGGGCGAGCTG	180
Db	584	GGCGAGGACAGGAGGGGCGAGGACAGGAGGGCGAGGAGGGGCGAGGAGGGGCGAGG	643
Qy	181	CTGCGGGAGAGGTGAACCTGCTGGAGCAGAGCTGACGAGCTGCGGGGCCAGGCCGCC	240
Db	644	CAGGAGGGGCGAGGACAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGG	703
Qy	241	CTGGGCGCGCAGATGGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCTCCAGCGCGGAG	300
Db	704	CAGGAGCAGGAGGGGCGAGGACAGGAGGGGCGAGGCGAGGAGGGGCGAGGAGGGG	763
Qy	301	CTGGAGCGGCTCGGGGCGAGCTGGGGAGGAGCGGCAAGGCCATGACCAATGTCCTCG	360
Db	764	CAGGAGGGGCGAGGAGGGGCGAGGACAGGAGGGGCGAGGAGGGGCGAGGAGGGG	823
Qy	361	GGCTCCAGCATGACGGCTCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	420
Db	824	CAGGAGGGGCGAGGACAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGG	883
Qy	421	AAACAGCTGCACGACAGCTAGCTGGCCATGTACACAGCGGAACCAAGCGCCTGGAGAAGGCC	480
Db	884	CAGGAGCAGGAGGGGCGAGGAGGGGCGAGGACAGGAGGGGCGAGGAGGGGCGAGGAGG	943
Qy	481	CTGCGAGCAGCTGGCACTGGGGACAGCGCCGGGAGGCCCTTGGAGGTTGACCTGGAAAGGG	540





XX	14-MAR-2002 (first entry)	
DT	Epstein-Barr virus nuclear antigen coding sequence.	
XX		
DE		
XX		
XX	Epstein barr virus; protein-protein interaction; fusion protein;	
KW	yeast two-hybrid; bioactive agent; therapeutic drug; ds.	
KW		
XX		
OS	Human herpesvirus type 4.	
XX		
PN	US6316223-B1.	
XX		
FD	13-NOV-2001.	
XX		
PF	22-JUL-1999; 99US-0359081.	
XX		
PR	30-MAR-1998; 98US-0050863.	
XX		
PA	(RIGE-) RIGEL PHARM INC.	
XX		
PI	Payan D, Luo Y, Huang B;	
XX		
DR	WPI; 2002-088877/12.	
XX		
CC	Detecting protein-protein interactions, useful screening candidate	
PT	bioactive agents or therapeutic drugs, comprises employing mammalian	
PT	cloning system cell comprising a bait vector and a test vector -	
XX		
PS	Disclosure; Column 27-30; 29pp; English.	
XX		
CC	The invention relates to the detection of an interaction between a bait	
CC	protein and a test protein, comprising employing a mammalian host cell	
CC	comprising a bait vector and a test vector, where the vectors contain	
CC	genes that encode proteins, and the method uses fusion proteins whose	
CC	expression and interaction with each other results in transcriptional	
CC	activation. The method and vectors are useful for analysing	
CC	protein-protein interactions between known proteins and for isolating,	
CC	cloning or characterising unknown proteins. The method is also useful	
CC	for screening candidate bioactive agents that modulate the	
CC	protein-protein interaction between a bait protein and a test protein	
CC	and thus are useful for the identification of novel therapeutic drugs.	
CC	The present method employs mammalian cloning systems that are highly	
CC	stable, and designed to reduce the background signals frequently found	
CC	in other systems such as the yeast two-hybrid system. The current	
CC	sequence represents the coding sequence for the Epstein-Barr virus	
CC	nuclear antigen, the viral replication protein, which, on binding to the	
CC	Epstein-Barr origin of replication (see AAI64274) on the bait vector,	
CC	allows vector replication to occur.	
XX		
SQ	Sequence 2580 BP; 632 A; 512 C; 1055 G; 381 T; 0 other;	
	Query Match 17.9%; Score 114.4; DB 24; Length 2580;	
	Best Local Similarity 49.8%; Pred. No. 6.8e-11;	
	Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;	
Qy	1 GAGGTGAACCCCAAGGCTAGCGAGATCCTCGGTCTCAAGGCACAGCTGAAGGACACGCGG 60	
Db	847 GAGGGCAGGACGAGGAGGAGGGCGAGGAGGGCGCAGGAGGGCGAGGAGGAGGAGGGG 906	
Qy	61 GGCAAGCTGGAGGGCCCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAG 120	
Db	907 CAGGAGCAGGAGGGGGCAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGGCGAGGAG 966	
Qy	121 GGCTTGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAGACGAGGCGGAGCTG 180	
Db	967 GGGCAGGAGCAGGAGGGGGCGAGGAGCAGGAGGGGGCAGGAGGGGGCAGGAGGAG 1026	
Qy	181 CTGCGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGGCCAGGCGCC 240	
Db	1027 CAGAGGGGCGAGGACGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGAGGAGGAGGG 1086	
Qy	241 CTGCGCCCGACATGGGGCGCCGCCACCTTCCCGAGGACGCTCCCTGCGCCCTGACGCGGAG 300	



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:27:06 ; Search time 37.4937 Seconds  
(without alignments)  
5226.648 Million cell updates

Title: US-09-513-888C-1 COPY 4912 5550

Perfect score: 639  
Sequence: 1 gaggtgaacgcccaaggctag.....tggaggagcagggtctctccg 639

Scoring table: IDENTITY NUC

Gapex 10.0, Gapext 1.0

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

[illegible]

```
Database : ISSUED_PATENTS_NA:
1: /cmt2 6/ptodata/3/ina/58 COMB 888: *
```

```
1: /cgnz_6/ptodata/z/ina/3A_COMB.seq:
2: /cgnz_6/ptodata/2/ina/5B_COMB.seq:*
```

3: /cqn2\_6/ptodata/2/ina/6A\_COMB.seq: \*

4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq: \*

5: /cgn2\_6/ptodata/2/ina/pCTUS\_COMB.se

6: /cgn2\_6/ptodata/2/ina/backfiles1.se

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Quarry	Length	ID	Description
1	114.4	17.9	1926	4	US-09-249-585A-2
2	114.4	17.9	2580	3	US-09-050-863-2
3	114.4	17.9	2580	4	US-09-359-081-2
4	114.4	17.9	5452	2	US-09-130-114-1
5	114.4	17.9	9600	4	US-08-910-647-1
6	114.4	17.9	9600	4	US-08-910-647-1
7	114.4	17.9	9600	4	US-09-620-925-1
8	114.4	17.9	10596	1	US-07-884-811-15
9	114.4	17.9	10596	1	US-07-885-971-15
10	114.4	17.9	10596	1	US-08-787-783A-15
11	114.4	17.9	10596	2	US-08-194-088B-15
12	114.4	17.9	10596	2	US-08-194-087-15
13	114.4	13.9	3489	2	US-08-728-323A-1
14	88.8	13.9	3489	2	US-09-298-568-1
15	88.8	13.9	32207	2	US-08-770-379-20
16	88.8	13.9	32207	4	US-08-757-669A-20
17	88.8	13.9	32207	4	US-09-230-371A-20
18	88.6	13.9	2793	1	US-08-209-747-1
19	88.6	13.9	2793	1	US-08-458-298-1
20	81.2	12.7	1926	4	US-09-249-585A-4
21	81.2	12.7	1931	2	US-09-130-114-2
22	81.2	12.7	16442	3	US-08-781-891-208
23	78.4	12.3	9551	1	US-08-056-200-93
24	78.4	12.3	9551	2	US-08-800-644-93
25	76	11.9	2338	1	US-08-425-069-1
26	76	11.9	2338	1	US-08-317-844B-1
27	75	11.7	1995	1	US-08-425-069-3
28	75	11.7	1995	1	US-08-425-069-3
29	75	11.7	1995	1	US-08-425-069-3
30	75	11.7	1995	1	US-08-425-069-3
31	75	11.7	1995	1	US-08-425-069-3
32	75	11.7	1995	1	US-08-425-069-3
33	75	11.7	1995	1	US-08-425-069-3
34	75	11.7	1995	1	US-08-425-069-3
35	75	11.7	1995	1	US-08-425-069-3
36	75	11.7	1995	1	US-08-425-069-3
37	75	11.7	1995	1	US-08-425-069-3
38	75	11.7	1995	1	US-08-425-069-3
39	75	11.7	1995	1	US-08-425-069-3
40	75	11.7	1995	1	US-08-425-069-3
41	75	11.7	1995	1	US-08-425-069-3
42	75	11.7	1995	1	US-08-425-069-3
43	75	11.7	1995	1	US-08-425-069-3
44	75	11.7	1995	1	US-08-425-069-3
45	75	11.7	1995	1	US-08-425-069-3
46	75	11.7	1995	1	US-08-425-069-3
47	75	11.7	1995	1	US-08-425-069-3
48	75	11.7	1995	1	US-08-425-069-3
49	75	11.7	1995	1	US-08-425-069-3
50	75	11.7	1995	1	US-08-425-069-3
51	75	11.7	1995	1	US-08-425-069-3
52	75	11.7	1995	1	US-08-425-069-3
53	75	11.7	1995	1	US-08-425-069-3
54	75	11.7	1995	1	US-08-425-069-3
55	75	11.7	1995	1	US-08-425-069-3
56	75	11.7	1995	1	US-08-425-069-3
57	75	11.7	1995	1	US-08-425-069-3
58	75	11.7	1995	1	US-08-425-069-3
59	75	11.7	1995	1	US-08-425-069-3
60	75	11.7	1995	1	US-08-425-069-3
61	75	11.7	1995	1	US-08-425-069-3
62	75	11.7	1995	1	US-08-425-069-3
63	75	11.7	1995	1	US-08-425-069-3
64	75	11.7	1995	1	US-08-425-



REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 949-8711

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2580 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-359-081-2

Query Match 17.9%; Score 114.4; DB 4; Length 2580;  
Best Local Similarity 49.8%; Pred. No. 4.7e-14;  
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 1 GAGGTGAACCCAGGCTAGCAGATCTTGGGTCTCAAGGCACACCTGAAGACACGCGG 60  
DB 847 GAGGGGACAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 906  
QY 61 GSCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCCACG 120  
DB 907 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 966  
QY 121 GGCCTGGAGCTGGAGGCTCTGTGAGATGAGCTGCAGCGCAAGAAAGAGCGGAGCTG 180  
DB 967 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1026  
QY 181 CTGCGGAGAGAGGTGAACCTGTGAGCAGAGCTGAGGAGCTGCGGGCCCGCCCGCC 240  
DB 1027 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1086  
QY 241 CTGGCCCGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCCCTGAGCGGGAG 300  
DB 1087 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1146  
QY 301 CTGAGCGGCTGCGGGCCCGAGCTGCGGAGGAGCGGCAAGGCCATGACCATGATGCTCT 360  
DB 1147 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1206  
QY 361 GCCTTCAGCATGAGCGGCTCGTGTGAAGAGGAGAGAGGAGAGGATGATTCAGTACCAG 420  
DB 1207 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1266  
QY 421 AAACAGCTGCACAGCTAGCTGCGCATGTACACCGGACCGACCGGCTCGAGAGGCC 480  
DB 1267 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1325  
QY 481 CTGACAGCTGGCACTGGGACAGCGCCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540  
DB 1326 --GGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1383  
QY 541 GCTGATCCCTTACAGAGACATCATAGCCATGATGATCTGAGGGGCTGCTTGGGAAGGC 600  
DB 1384 GTGAGGGCCGGGCTCGAGGAGGTAGTGGAGGCGGGGTGAGGAGGAGGAGGAGGAGG 1443  
QY 601 GAGTCTGGGACCTGCGACTGGGAGCGAGGCGTCTC 636  
DB 1444 GGGGTAGAGGACGTGAAGAGCCAGGCGGGGGAAGTC 1479

RESULT 4

US-09-130-114-1/c

Sequence 1, Application US/09130114

Patent No. 5976807

GENERAL INFORMATION:

APPLICANT: Horlick, Robert A.

APPLICANT: Damaj, Basam B.

APPLICANT: Robbins, Alan K.

TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes

FROM MULTIPLE TRANSFECTED EPISOMES

FILE REFERENCE: 0867/1D903US1

CURRENT APPLICATION NUMBER: US/09/130.114

CURRENT FILING DATE: 1998-08-06

NUMBER OF SEQ ID NOS: 36

SOFTWARE: PastSeq for Windows Version 3.0

SEQ ID NO 1

LENGTH: 5452

TYPE: DNA

ORGANISM: VEBNA

US-09-130-114-1

Query Match 17.9%; Score 114.4; DB 2; Length 5452;  
Best Local Similarity 49.8%; Pred. No. 5.2e-14;  
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 1 GAGGTGAACCCAGGCTAGCAGATCTTGGGTCTCAAGGCACACCTGAAGACACGCGG 60  
DB 1958 GAGGGGACAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1899  
QY 61 GSCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCCACG 120  
DB 1898 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1839  
QY 121 GGCCTGGAGCTGGAGGCTCTGTGAGATGAGCTGCAGCGCAAGAAAGAGCGGAGCTG 180  
DB 1838 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1779  
QY 181 CTGCGGAGAGAGGTGAACCTGTGAGCAGGAGCTGAGGAGCTGCGGGCCCGCCCGCC 240  
DB 1778 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1719  
QY 241 CTGGCCCGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCCCTGAGCGGGAG 300  
DB 1718 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1659  
QY 301 CTGAGCGGCTGCGGGCCCGAGCTGCGGAGGAGCGGCAAGGCCATGACCATGATGCTCT 360  
DB 1658 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1599  
QY 361 GCCTTCAGCATGAGCGGCTCGTGTGAAGAGGAGAGAGGAGAGGATGATTCAGTACCAG 420  
DB 1598 CAGGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1539  
QY 421 AAACAGCTGCACAGCTAGCTGCGCATGTACACCGGACCGACCGGCTTGGAGAGGCC 480  
DB 1538 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1480  
QY 481 CTGACAGCTGGCACTGGGACAGCGCCGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540  
DB 1479 --GGAGGGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1422  
QY 541 GCTGACATCCCTTACAGAGACATCATAGCCATGATGATCTGAGGGGCTGCTTGGGAAGGC 600  
DB 1421 GTGAGGGCCGGGCTCGAGGAGGTAGTGGAGGCGGGGTGAGGAGGAGGAGGAGGAGG 636  
QY 601 GAGTCTGGGACCTGCGACTGGGAGCGAGGCGTCTC 636  
DB 1361 GGGGTAGAGGACGTGAAGAGCCAGGCGGGGGAAGTC 1326

RESULT 5

US-08-910-647-1

Sequence 1, Application US/08910647

Patent No. 6251433

GENERAL INFORMATION:

APPLICANT: Zuckermann et al.

TITLE OF INVENTION: Compositions and Methods for

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

```
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/910,647
  FILING DATE:
  CLASSIFICATION: 514
  ATTORNEY/AGENT INFORMATION:
    NAME: Fujita, Sharon M.
    REGISTRATION NUMBER: 38,459
    REFERENCE/DOCKET NUMBER: 1218.002
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (510) 923-2706
    TELEFAX: (510) 655-3542
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 9600 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

Query Match      17.9%; Score 114.4; DB 4; Length 9600;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACCCAGAGCTAGCAGATCTCTGAGTCTCAAGGCACAGCTGAAGACACGCGG 60
DB 893 GAGGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 952
QY 61 GCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 993 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1012
QY 121 GCGTGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 1013 GCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1072
QY 181 CTGCGGAGAGAGTGAACCTGCTGAGCAGAGAGCTGAGAGAGCTGCGGGGCCAGGCCGCC 240
DB 1073 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1132
QY 241 CTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCAGCGGAG 300
DB 1133 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1192
QY 301 CTGGAGCGGCTCGGGCCCGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
DB 1193 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1252
QY 361 GCGTCCAGCATAGCGGCTGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 1253 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1312
QY 421 AAACAGCTCAGCAGAGCTACCTGCGCATGTACAGCGGAGGAGGAGGAGGAGGAGGAGG 480
DB 1313 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1371
QY 481 CTGCAGCAGCTGCGCAGTGGGAGCAGCGCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
DB 1372 --GGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1429
QY 541 GCTGACATCCCTACGAGACATCATAGCCATGAGATCTGAGGAGGAGGAGGAGGAGGAGG 600
DB 1430 GTGGAGGGCGGGGTCTGAGAGGAGTGTGGAGGCGGGGTCTGAGAGGAGTGTGGAGGCG 1489
QY 601 GAGTCTGGGAGCTGTCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
```

DB 1490 GGGGTAGAGGACGTGAAGAGCCAGGGGGGAAGTC 1525

## RESULT 6

```
US-09-620-925-1
Sequence 1, Application US/09620925
Patent No. 6468986
GENERAL INFORMATION:
  APPLICANT: Zuckermann et al.
  TITLE OF INVENTION: Compositions and Methods for
    Polynucleotide Delivery
  NUMBER OF SEQUENCES: 4
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Chiron Corporation
    STREET: 4560 Horton Street
    CITY: Emeryville
    STATE: California
    COUNTRY: U.S.A.
    ZIP: 94608-2916
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/620,925
    FILING DATE: 21-Jul-2000
    CLASSIFICATION: <Unknown>
    PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 08/910,647
      FILING DATE: <Unknown>
    ATTORNEY/AGENT INFORMATION:
      NAME: Fujita, Sharon M.
      REGISTRATION NUMBER: 38,459
      REFERENCE/DOCKET NUMBER: 1218.002
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (510) 923-2706
      TELEFAX: (510) 655-3542
    INFORMATION FOR SEQ ID NO: 1:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 9600 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-620-925-1
```

```
Query Match      17.9%; Score 114.4; DB 4; Length 9600;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 1 GAGGTGAACCCAGAGCTAGCAGATCTCTGAGTCTCAAGGCACAGCTGAAGACACGCGG 60
DB 893 GAGGGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 952
QY 61 GCGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 993 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1012
QY 121 GCGTGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 1013 GCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1072
QY 181 CTGCGGAGAGAGTGAACCTGCTGAGCAGAGAGCTGAGAGAGCTGCGGGGCCAGGCCGCC 240
DB 1073 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1132
QY 241 CTGGCCCGGACATGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCGCAGCGGAG 300
DB 1133 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1192
```



```

; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
US-07-885-971-15

Query Match 17.9%; Score 114.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

```

```

QY 1 GAGGTGAACGCCAAGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
DB 2427 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2486
QY 61 GCGAAGCTGGAGGGCTGAGAGTGAAGGACCCAGGACCTGGAGGGCGCCCTGCGCAACCAAG 120
DB 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2546
QY 121 GGCCTGGAGCTGGAGGTCTGTAGAAATGAGTGCAGCGCAAGAAAGAGGCGGAGCTG 180
DB 2547 GCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2606
QY 181 CTGCGGGAAGGTGAACCTCTGTGAGCAGGAGCTGAGGAGCTGGGGCCCGCCAGCCGCC 240
DB 2607 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2666
QY 241 CTGGCCCGGACATGGGGCGCCCGCCCTTCCCGAGGAGCTCCCTGCGCTGAGCGGAG 300
DB 2667 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2726
QY 301 CTGGAGCGGCTCGGGCCGAGCTGCGGAGGAGCGGCAAGGCGCATGACAGATGCTCTCG 360
DB 2727 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2786
QY 361 GGCCTCCAGCATGAGCGGCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 2787 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2846
QY 421 AAACAGCTGCACAGAGTACGTGGCCATGTACCAAGCGGAAACAGCGCTGGAGAAGGCC 480
DB 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2905
QY 481 CTGAGCAGCTGGCAGCTGGGACACAGCCCGGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540
DB 2906 --GGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2963
QY 541 GCTGACATCCCTAGCAGGACATCATAGCCACTGAGATCTGAGGAGGCTGCTGGGAAGGC 600
DB 2964 GTGGAGGCGGGGTCAAGAGAGTAGTGGAGGCGGGGTGAGAGGCTGAGAGGAGGAGGAG 3023
QY 601 GAGTCTGGGAGCTGGCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
DB 3024 GGGGTAGAGGACGTGAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3059

```

## RESULT 9

```

US-08-087-783A-15
; Sequence 15, Application US/08087783A
; Patent No. 5547856
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

```

```

; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,783A
; FILING DATE: 13-Jul-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0755779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-087-783A-15

```

```

Query Match 17.9%; Score 114.4; DB 1; Length 10596;
Best Local Similarity 49.8%; Pred. No. 5.6e-14;
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;
QY 1 GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGG 60
DB 2427 GAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2486
QY 61 GCGAAGCTGGAGGGCTGAGAGTGAAGGACCCAGGACCTGGAGGGCGCCCTGCGCAACCAAG 120
DB 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2546
QY 121 GGCCTGGAGCTGGAGGTCTGTAGAAATGAGTGCAGCGCAAGAAAGAGGCGGAGCTG 180
DB 2547 GCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2606
QY 181 CTGCGGGAAGGTGAACCTCTGTGAGCAGGAGCTGAGGAGCTGGGGCCCGCCAGCCGCC 240
DB 2607 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2666
QY 241 CTGGCCCGGACATGGGGCGCCCGCCCTTCCCGAGGAGCTCCCTGCGCTGAGCGGAG 300
DB 2667 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2726
QY 301 CTGGAGCGGCTCGGGCCGAGCTGCGGAGGAGCGGCAAGGCGCATGACAGATGCTCTCG 360
DB 2727 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2786
QY 361 GGCCTCCAGCATGAGCGGCTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 2787 CAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2846
QY 421 AAACAGCTGCACAGAGTACGTGGCCATGTACCAAGCGGAAACAGCGCTGGAGAAGGCC 480
DB 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2905
QY 481 CTGAGCAGCTGGCAGCTGGGACACAGCCCGGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540

```



Db 2906 --GGAGGGGACAGGACAGGAGGGGACAGGAGGAGGTGGAGCGCGGTCTCGAGGAGCA 2963  
QY 541 GCTGACATCCCTACGAGGACATACGACCTGAGATCTGAGGGGCTGCTTGGGAAGGC 600  
Db 2964 GTGGAGGCGGGGTCTGAGAGGTAGTGGAGGCGGGGTCTGAGGAGGTAGTGGAGGCGGCC 3023  
QY 601 GAGTCTGGGACCTGGCACTGGGAGGCGAGGCTCTC 636  
Db 3024 GGGGTAGAGGACGTGAAGAGCCAGGCGGGGGAAGTC 3059

## RESULT 10

US-08-194-088B-15  
; Sequence 15, Application US/08194088B  
; Patent No. 5580963  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.  
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/194,088B  
; FILING DATE: 09-FEB-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/884811  
; FILING DATE: 18-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gallegos, R. Thomas  
; REGISTRATION NUMBER: 32,692  
; REFERENCE/DOCKET NUMBER: 755D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-2614  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10596 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-194-088B-15  
Query Match 17.9%; Score 114.4; DB 1; Length 10596;  
Best Local Similarity 49.8%; Pred. No. 5.6e-14;  
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 1 GAGGTGAACGCAAGCTTAGCGAGATCTGGTCTCAAGGCACAGCTGAAGGACACGGCG 60  
Db 2427 GAGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2486  
QY 61 GGCAGAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCTTGGCCACCAAG 120  
Db 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2546  
QY 121 GGCCTGGAGCTGAGGTCTGTGAGAAATGAGTGCAGCGCAAGAAACGAGCGGAGCTG 180  
Db 2547 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2606  
QY 181 CTGCGGGAGAGGTGAACCTGCTGAGCAGGAGGAGTGCAGGAGCTCGGGGCCAGCCGCC 240

Db 2607 CAGGAGGGGACGAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2666  
QY 241 CTGGCCCCGACATAGGGGCCCCACACCTTCCCCAGAGGACGTCCTTCCCTGCGAGCGGAG 300  
Db 2667 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2726  
QY 301 CTGAGCGGCTGCGGGCCGAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360  
Db 2727 CAGGAGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2786  
QY 361 GCGTTCCAGCATGAGCGGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420  
Db 2787 CAGGAGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2846  
QY 421 AACAGCTGCAGCAGAGCTAGCTGGCCATGTACAGCGGAACACCGCTTGAGAGAGGCC 480  
Db 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2905  
QY 481 CTGAGCAGCTGGGACGCTGGGGACAGCCCGGGGAGCCCTTGGAGGTTGACCTGGAAGGG 540  
Db 2906 --GGAGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2963  
QY 541 GCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGCTGCCCTGGGAAGGC 600  
Db 2964 GTGGAGGCGGGGTCTGAGGAGGTAGTGGAGGCGGGGTCTGAGGAGGAGGAGGAGG 3023  
QY 601 GAGTCTGGGACCTGGCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636  
Db 3024 GGGGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTC 3059

## RESULT 11

US-08-194-087-15  
; Sequence 15, Application US/08194087  
; Patent No. 5879910  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.  
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/194,087  
; FILING DATE: 18-MAY-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Girger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: 779  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10596 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

US-08-194-087-15

Query Match 17.9%; Score 114.4; DB 2; Length 10596;  
Best Local Similarity 49.8%; Pred. No. 5.6e-14;  
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACGCGG 60  
Db 2427 GAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2486  
QY 61 GGCAGCTGAGGAGGCTGAGAGTGAAGACCCAGGACCTGAGAGGCGCCCTGCGACCAAG 120  
Db 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2546  
QY 121 GGCCTGAGCTGAGGCTGTGAGATGAGCTGACGCGCAAGAAACGAGCGGAGCTG 180  
Db 2547 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2606  
QY 181 CTCGGGAGAGGTGAACCTGCTGAGCAGGAGCTGACAGGAGCTGCGGCGCCAGCGCC 240  
Db 2607 CAGGAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2666  
QY 241 CTGGCCCGGACATGCGGCGCCACCTTCCCGAGGAGCTCCCTGCGCCCTGCGCGGAG 300  
Db 2667 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2726  
QY 301 CTGGAGCGGCTCGGCGCGAGCTGCGGAGGAGCGGAGGAGGAGGAGGAGGAGGAGG 360  
Db 2727 CAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2786  
QY 361 GGCCTCCAGCATGAGCGGCTCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
Db 2787 CAGGAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2846  
QY 421 AACAGCTGACAGAGCTACCTGCGCATGTACAGCGGAAACCGGCTGAGAGGCGCC 480  
Db 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2905  
QY 481 CTGACGAGCTGCGAGCTGGGACACGCGCGGAGCGCCCTTGAGAGGTTGACCTGAAGG 540  
Db 2906 --GGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2963  
QY 541 GCTGACATCCCCACGAGACATCATAGCCTGAGACTGAGGGGCTGCTTGGGAAGGC 600  
Db 2964 GTGAGGCGCGGCTCGAGAGGTAGTGGAGCGCGGGTCTGAGAGGTAGTGGAGCGCGC 3023  
QY 601 GAGTCTGGGAGCTGCGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 636  
Db 3024 GGGGTAGAGGAGCTGAAGAGACCGAGGGGGGAAGTC 3059

RESULT 12  
PCT-US93-04648-15  
; Sequence 15, Application PC/TUS9304648  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Natalie A., Mark, Melanie R  
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/04648  
; FILING DATE: 19930517  
; CLASSIFICATION:

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/884811  
; FILING DATE: 18-MAY-92  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/885971  
; FILING DATE: 18-MAY-92  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: 755,779P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/952-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10596 bases  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; PCT-US93-04648-15

Query Match 17.9%; Score 114.4; DB 5; Length 10596;  
Best Local Similarity 49.8%; Pred. No. 5.6e-14;  
Matches 317; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 1 GAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACGCGG 60  
Db 2427 GAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2486  
QY 61 GGCAGCTGAGGAGGCTTGGAGCTGAGGACCCAGGACCTGAGAGGCGCCCTGCGACCAAG 120  
Db 2487 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2546  
QY 121 GGCCTGAGCTGAGGCTGTGAGATGAGCTGACGCGCAAGAAACGAGCGGAGCTG 180  
Db 2547 GGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2606  
QY 181 CTGGCGGAGAGGTGAACCTGCTGAGCAGGAGCTGCGGAGGAGCTGCGGCGCCAGCGCC 240  
Db 2607 CAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2666  
QY 241 CTGGCGCGCATATGGGCGCGCCACCTTCTCCCGAGGAGCTCCCTGCGCCCTGCGCGGAG 300  
Db 2667 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2726  
QY 301 CTGAGAGCGGCTGCGGCGCGAGCTGCGGAGGAGCGGCAAGCGCCATGACAGATGCTCTG 360  
Db 2727 CAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2786  
QY 361 GGCCTCCAGCATGAGCGGCTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
Db 2787 CAGGAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2846  
QY 421 AACAGCTGACAGAGCTAGCTGGCCATGTACCGCGGAAACCGAGCGCTTGAGAGGCGCC 480  
Db 2847 CAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2905  
QY 481 CTGAGAGCTGCGACGCTGGGAGACAGCGCGGAGGAGCGCCCTTGGAGGTTGACCTGAAGG 540  
Db 2906 --GGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2963  
QY 541 GCTGACATCCCCACGAGGACATCATAGCCTGAGACTGAGGGGCTGCTTGGGAAGGC 600  
Db 2964 GTGAGGCGCGGCTCGAGAGGTAGTGGAGCGCGGGTCTGAGAGGTAGTGGAGCGCGC 3023  
QY 601 GAGTCTGGGAGCTGCGACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 636  
Db 3024 GGGGTAGAGGAGCTGAAGAGACCGAGGGGGGAAGTC 3059

RESULT 13  
US-08-728-323A-1

Sequence 1, Application US/08728323A  
Patent No. 5948676  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
TITLE OF INVENTION: Encoding Same And Uses Thereof  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/728,323A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSK/SKS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3489 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3489  
US-08-728-323A-1

Query Match 13.9%; Score 88.8; DB 2; Length 3489;  
Best Local Similarity 50.5%; Pred. No. 4.3e-09;  
Matches 216; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 65 AGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCC 124  
DB 1802 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1861

QY 125 TGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGAGCGCAAGAAAGAGCGGAGCTGCTGC 184  
DB 1862 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1921

QY 185 GGGAGAGGTGAACCTGCTGAGCAGCAGGAGCTGCGAGGAGCTGGGGCCCGCCCTGG 244  
DB 1922 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1981

QY 245 CCCGCGACATGGGCGCGCCACCTTTCCCGAGGAGCTCCCTGCCCTGCGAGCGGAGCTGG 304  
DB 1982 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2041

QY 305 AGCGGCTCGGGCCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACACAGATGTCCTCGGGT 364  
DB 2042 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2101

QY 365 TCCAGCATGACCGGCTGTGTGGAAGGAGGAGAGAGGAGGATGATTCAGTACCGAAGC 424  
DB 2102 AGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2161

QY 425 AGCTGACAGCAGCTACGTGGCCATGTACACAGCGAACCAGCGCTGCGAGAGGCCCTGC 484  
DB 2162 ATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGC 2221

QY 485 AGCAGCTG 492  
DB 2222 AGCAGCAG 2229

RESULT 14  
US-09-298-568-1  
Sequence 1, Application US/09298568  
Patent No. 6322792  
GENERAL INFORMATION:  
APPLICANT: Kieff, Elliott D.  
APPLICANT: Ballestas, Mary E.  
APPLICANT: Kaye, Kenneth M.  
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
FILE REFERENCE: 16412-10001R  
CURRENT APPLICATION NUMBER: US/09/298,568  
CURRENT FILING DATE: 1999-04-21  
EARLIER APPLICATION NUMBER: US 60/109,422  
EARLIER FILING DATE: 1998-11-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 3489  
TYPE: DNA  
ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-298-568-1

Query Match 13.9%; Score 88.8; DB 4; Length 3489;  
Best Local Similarity 50.5%; Pred. No. 4.3e-09;  
Matches 216; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 65 AGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCC 124  
DB 1802 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1861

QY 125 TGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGAGCGCAAGAAAGAGCGGAGCTGCTGC 184  
DB 1862 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1921

QY 185 GGGAGAGGTGAACCTGCTGAGCAGCAGGAGCTGCGAGGAGCTGGGGCCCGCCCTGG 244  
DB 1922 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 1981

QY 245 CCCGCGACATGGGCGCGCCACCTTTCCCGAGGAGCTCCCTGCCCTGCGAGCGGAGCTGG 304  
DB 1982 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2041

QY 305 AGCGGCTCGGGCCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACACAGATGTCCTCGGGT 364  
DB 2042 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2101

QY 365 TCCAGCATGACCGGCTGTGTGGAAGGAGGAGAGGAGGATGATTCAGTACCGAAGC 424  
DB 2102 AGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGCAGC 2161

QY 425 AGCTGACAGCAGCTACGTGGCCATGTACACAGCGAACCAGCGCTGCGAGAGGCCCTGC 484  
DB 2162 ATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGCAGCAGGATGAGC 2221

QY 485 AGCAGCTG 492  
DB 2222 AGCAGCAG 2229

RESULT 15  
US-08-770-379-20/c

Sequence 20, Application US/08770379  
Patent No. 5849584  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/770,379  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double.  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 13.9%; Score 88.8; DB 2; Length 32207;  
Best Local Similarity 50.5%; Pred. No. 5.7e-09;  
Matches 216; Conservative 0; Mismatches 212; Indels 0; Gaps 0;  
QY 65 AGCTGAGGGCCTGGAGCTGAGACCCAGGACCTGGAGGGCGCCCTTGGCGCACCAAGGGCC 124  
DB 20195 AGCAGGATGAGCAGCAGCAGGATGACGACGACGAGGATGAGCAGCAGCAGGATGAGCAGC 20136  
QY 125 TGGAGCTGAGGTCTGTGAGATGAGCTCAGCGCAAGAAACGAGGCGGAGCTGCTGC 184  
DB 20135 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 20076  
QY 185 GGGAGAGGTGACCTGCTGGACGAGAGCTCAGCAGCTCGGGCCCGCCCTGG 244  
DB 20075 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 20016  
QY 245 CCGCGACATGGGGCGGCCACCTTCCCGAGGACGTCCTCCCTGCGAGCGGGAGCTGG 304  
DB 20015 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 19956  
QY 305 AGCGGCTGCGGGCCGAGCTCGGGAGGAGCGGCAAGCCATGACCAAGATGTCTCTCGGGCT 364  
DB 19955 AGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGC 19896  
QY 365 TCCAGCATGCGGCTCGTGTGGAGGAGGAGGAGAGGTGATTCTAGTACCAGAAC 424  
DB 19895 AGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGG 19836  
QY 425 AGCTGACGAGAGCTACGTGGCCATGTACCGCGGAACCGAGCGCTTGGAGAGGCCCTGC 484

Db 19835 ATGAGCAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGCAGGATGAGCAGCAGGATGAGC 19776  
QY 485 AGCAGCTG 492  
Db 19775 AGCAGCAG 19768  
Search completed: June 15, 2003, 03:45:48  
Job time : 39.4937 secs



Db 279 GAGCTGTGCAGAAAGGCGAGAGCTGTGTGCTCTCGGGTGGCGCTGCGGAGGCGCGT 338  
Qy 61 GGCAAGCTGGAGGGGCTGAGCTGAGNACCAGAGCTTGGAGGGGCGCTCTGCGCACCAAG 120  
Db 339 GCTACGCTGCGGGGTGAGTGGAGGGCGGTGCGGGGGTCTACAGGAGGCGCGCGAGCTCGG 398  
Qy 121 GGCCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGAGCGCAAGAAAGAGAGGCGGAGCTG 180  
Db 399 GAGCTGGAGCTGGAGGCTGTCCAGAGAGCTGAGCGACACCGCCAGGAGCTGAGCAG 458  
Qy 181 CTGCGGGAGAAAGGTGAACCTCTGTGAGCAGAGCTGACAGAGCTGCGGGC----- 230  
Db 459 CTGCGGGAGAAAGCTGGGCGAGTTGATGCTGAGGCGCGCGACTCGGGAGCCCGCTGTG 518  
Qy 231 -----CCAGCGCGCCCTGGCGCGGAGCATGGGCGCGCC 264  
Db 519 CCACCTGTCACCGCTGACCCATTCTCTGCGCAGAGAGTGATGAGGCGCAAGTGCAGCG 578  
Qy 265 ACCTTCCCGGAGGAGCTCCCTGCGCTGCGAGCGGAGCTGGAGCGGCTGCGGGCGGAGCTG 324  
Db 579 CGAGCAGCGGGGTGGGGGCGAGCTTGGCGGCCAGGTGGAGCGATTGCGGGTGGAGCTG 638  
Qy 325 CGGAGGAGCGCGCAAGGCGCATGACAGATGCTCTGGGCTTCCAGCATGAGCGGCTCGTG 384  
Db 639 CAGCGGAGCGCGCGGGGTGAGGAGCAGCGGACAGCTTTGAGGGGAGCGGCTGGCC 698  
Qy 385 TGGAGGAGGAGAGGAGGAAGTGATTCAGTACAGAAACAGCTGCGACGAGCTACGTG 444  
Db 699 TGGCAGGCGAGAAAGGAGCAGGTATCCGCTACCAGAGAGCAGCTGCGACCAACTACATC 758  
Qy 445 GCCATGTACACGCGGAACAGCGCTGAGAGGCGCTGCGAGCGCTGCGAGCTGGCGAGC 504  
Db 759 CAGATGTACCGCGCGAACCGAGCTAGACAGAGCTGCGAGAGCTGAGCTGAGCTG 818  
Qy 505 AGCGCGGGGAGCGCTTGGAGGTTGACCTG-----GAAGGGGCTGACATCCCTACGAGG 559  
Db 819 GAGGCGCGGAGCTGCTGACTGGGCTGCGCGAGCAGCGCCCTGCACTGCTGCGTGGAGG 878  
Qy 560 ACATCATAGCCACTGAGATCTGAGGGGCTGCTGGAGGCGGAGCTGCGGAGCTGGCAC 619  
Db 879 AGATCATGCTACTGAGATCTAGGGCCCTCAGCAACAGAGCTCTGTAGGGAGCTCTGCCAG 938  
Qy 620 TGGG 623  
Db 939 AGG 942

RESULT 2  
US-10-171-581-63  
; Sequence 63, Application US/10171581  
; Publication No. US20030104426A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: Lineley, Peter  
; APPLICANT: Mao Mao  
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia  
; FILE REFERENCE: 9301-157-999  
; CURRENT APPLICATION NUMBER: US/10/171,581  
; CURRENT FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/298,914  
; PRIOR FILING DATE: 2001-06-18  
; NUMBER OF SEQ ID NOS: 366  
; SEQ ID NO 63  
; LENGTH: 5257  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: AB011124  
; DATABASE ENTRY DATE: 2001-06-18  
US-10-171-581-63

Query Match 19.6%; Score 125.2; DB 9; Length 5257;  
Best Local Similarity 70.9%; Pred. No. 3.8e-20;

Matches 166; Conservative 0; Mismatches 68; Indels 0; Gaps 0;  
Qy 286 GCCTTCAGCGGAGCTGGAGCGGCTGCGGCGGAGCTGCGGAGGAGCGCAAGGCCAT 345  
Db 3122 GCCTTCAGCGGAGGAGTGGGGCGGCTGACAGCCGAGCTGCGGCTGAGCGCGGCGCGG 3181  
Qy 346 GACCAGATGCTCTCGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGGAGAAG 405  
Db 3182 GAGGCCAGGGTGCCAGCTTCCGCGAGGAGCGCGCTGTGTGGTAGGAGAGAGAGAG 3241  
Qy 406 GTGATTCAGTACCAGAAACAGCTGCGACGAGCTACCTGGCCATGTACACGCGGAACAG 465  
Db 3242 GTGATCGAGTACCAGAAACAGCTGCGAGCTACGTGAGATGTACACGCGCAACAG 3301  
Qy 466 CGCTTGAGAGAGGCCCTGCGAGCAGCTGCGAGTGGGACAGCGCGGGGAGGCC 519  
Db 3302 CAGCTGAGGCGAGGCTGCGGAGCGCGGCGCGGCGCGGAGGAGGAGGAGGAGGAG 3355

RESULT 3  
US-09-962-832-225  
; Sequence 225, Application US/09962832  
; Patent No. US20020110821A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-74  
; CURRENT APPLICATION NUMBER: US/09/962,832  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,077  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,280  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 259  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 225  
; LENGTH: 2108  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-962-832-225

Query Match 12.4%; Score 79; DB 10; Length 2108;  
Best Local Similarity 49.1%; Pred. No. 1.8e-09;  
Matches 239; Conservative 0; Mismatches 245; Indels 3; Gaps 1;  
Qy 44 AGCTGAAGGACACGCGGGGCAAGCTGAGGGCTTGGAGCTGAGAGCTGAGAGCTGAGAGG 103  
Db 898 AGGGGAGCTGAAGCACCTGGATCAGCAGGAGAGAGGAGGAGGAGCTCCAGAGCAGAGA 957  
Qy 104 GCGCCCTGCGCACCAAGGCTTGGAGCTTGGAGTCTGTGAGATGAGCTGAGCGCAGA 163  
Db 958 TGGGGAGCTGAAGCACCTTGGAGCAGGAGGAGGGGAGCCCTAAGCATCTGAGAGCAGG 1017  
Qy 164 AGAACAGAGGCGGAGCTGCTGCGGAGAGAGGTGAACCTGCTGGAGCAGGAGCTGAGGAGC 223  
Db 1018 AGGGGCACTTGGAGCAGCTGAGGAGCAGGAGGAGGAGCTGAGACCTTGGAGCAGCAGG 1077  
Qy 224 TCGGGGCGCCAGGCGCGCTTGGCGCGGAGCATGCGGGCGCGCCACCTTCCCGAGAGCGTCC 283  
Db 1078 AGGGGAGCTGGAGCACCTTGGAGCACCGAGGAGGAGGAGCTGGGGCTCCAGAGCAGCAGG 1137  
Qy 284 CTGCCCTGCGAGGAGCTGGAGCGGCTGCGGGCGGAGCTGCGGGAGGAGCGGCAAGGCC 343  
Db 1138 TGCTGAGCTGAAGCAGCTAGAGAAAGCAGCAGGCGGAGCCCAAGCACTTGGAGAGGAGG 1197  
Qy 344 ATGACAGATGCTCTCGGGCTTCCAGCATGAGCGGCTGTGTGGAAGGAGGAGAGAGAGA 403  
Db 1198 AGGGGAGCTGAAGCACCTTGGTGCAGCAGGAGGAGGAGCTGAAGCATCTGTGTGAGCAGG 1257  
Qy 404 AGGTGATTGAGTACCAAGAAACAGCTGCGACGAGCTACGTGGGCCATGTACAGCGGAGACC 463  
Db 1258 AGGGGAGCTGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1317



APPLICANT: Secrist, Heather  
APPLICANT: Indirias, Carol Yoseph  
APPLICANT: Benson, Darin R.  
APPLICANT: Elliot, Mark  
APPLICANT: Mannion, Jane  
APPLICANT: Kalos, Michael D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C9  
CURRENT APPLICATION NUMBER: US/09/738,973  
CURRENT FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 587  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 337  
LENGTH: 422  
TYPE: DNA  
ORGANISM: Homo sapiens  
NAME/KEY: misc feature  
LOCATION: (1)-(422)  
OTHER INFORMATION: n = A,T,C or G  
US-09-738-973-337

Query Match 12.1%; Score 77.4; DB 10; Length 422;  
Best Local Similarity 49.6%; Pred. No. 3.8e-09;  
Matches 198; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

Qy 143 AGATGAGCTGAGCGGCAAGAAAGACGAGCGGAGCTCTCTGGGAGAGGTGAACCTGC 202  
Db 6 AGGAACAGTTACAGCAGCAGCAGCAACAGCAGCTGTTGCAACAGCAGCAGCAAAATTGC 65  
Qy 203 TGGAGCAGGAGCTGAGGAGCTGGGGCCAGGCGCCCTGGCCCGACATGGGGCGC 262  
Db 66 AGCAGCAACAATGAGCTCTCTCCCTGGAGCCGAGGAGGAGGAGGTGAGCTGG 125  
Qy 263 CCACCTTCCCGAGGAGCTCTCTGCGAGCGGAGCTGGAGCGGCTGGGGCCGAGC 322  
Db 126 AGCTCATCGCGTGAACCTGGGCTCAGAGCAGGAGCTGGAGCAGCGGCGAGGTTGG 185  
Qy 323 TGGCGGAGAGCGGCAAGCCATGACAGATGTCTCTGGGTTCCAGCATGAGCGGCTGC 382  
Db 186 AGCGCAGCAGGAGCTGGAACGCGCAGCAGGAGCAGCGCAGCTGAGCTCAAACTGCAGG 245  
Qy 383 TGTGAAGGAGGAGGAGGAGGAGTTCAGTACGAGAAACAGCTGAGCAGAGCTACG 442  
Db 246 AGGAGCTGAGAGCTGGAGCAACAGCTGGAGCAGCAGCAGCAGCTGGAGCAGCAGG 305  
Qy 443 TGGCCATGTACCGCGGAACAGCGCCCTGGAGAGGCGCTGCGAGCAGCTGGCACTGGGG 502  
Db 306 AGGTGAGCTGAGAGCTGACCCCGGTGGAGCTAGGCGCCAGCAGCAGGAGGTGCGAGCTGG 365  
Qy 503 ACAGCGCGGGAGCGCTTGGAGTTGACCTGGAGAGGGG 541  
Db 366 AGCTGACCCCGCTGCGAGCGGAGCTGCGAGCTGGAAGTGG 404

## RESULT 7

US-09-764-891-6987/c  
Sequence 6987, Application US/09764891  
Publication No. US20030077808A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC006  
CURRENT APPLICATION NUMBER: US/09/764,891  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 10231  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 6987  
LENGTH: 22595  
TYPE: DNA  
ORGANISM: Homo sapiens

## FEATURE:

NAME/KEY: SITE  
LOCATION: (9701)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-891-6987

Query Match 12.1%; Score 77.2; DB 9; Length 22585;

Best Local Similarity 46.9%; Pred. No. 5.2e-09;  
Matches 241; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

Qy 13 AAGCTAGCAGATCTCTGGGTCTCAAGGCACAGCTGAAGACACGCGGGGCAAGCTGGAG 72  
Db 10356 AAGGAGATGAGATGACAGGAGGGGACGACATGACAGAGGGGACGACAGGTGACAGAGGG 10297  
Qy 73 GGCTGAGCTGAGGACCCAGGAGCTGGAGGGGCCCTGCGCACCAGAGGGGCTGGAGCTG 132  
Db 10296 ACGCAGATGACAGGAGGGGACGGGCTGACAGAGGAGATGACAGGAGGGGCTGACAGTACAG 10237  
Qy 133 GAGTCTGTGAGAAATGAGCTGCAGCGCAAGAAACAGCGGAGCTGTCTCGGGGAGAAG 192  
Db 10236 GAGGGTTGGGCTGACAGGAGGGGACACAGATGACAGAGGGGCGCAGATGACAGAGCGG 10177  
Qy 193 GTGAACCTGCTGGAGCAGGAGCTGACAGAGCTGCGGCCCCAGCGCCCTTGGCCCGGAC 252  
Db 10176 ACGGGGCTGCAGAAAGAGATGACAGATGACAGAGGGGATGACAGATGACAGAGGAGATGACAG 10117  
Qy 253 ATGGGGCGCCACCTTCCCGGAGGAGCTCCCTGCTCCCTGAGCGGGAGCTGGAGCGGCTG 312  
Db 10116 ATGACAGAGGGGATGACAGGTGACAGGAGGGGACGCGGCTGACAGAGGGGCGGGCTGGAG 10057  
Qy 313 CGGGCCGAGCTGCGGGAGGAGCGGCAAGGCGCATGACAGATGTCTCGGGCTTCCAGCAT 372  
Db 10056 GGNACGGGGCTACAGGAGGAGATGACAGATGACAGAGGGGCTGACAGTACAGGATGGTCG 9997  
Qy 373 GAGCGCTCTGTGGAAGGAGGAGAGAGAGTATTAGTACCAAGAAACAGCTGCAG 432  
Db 9996 GGGCTGACAGAGGGGACGACATGACAGAGCGCTGCGGGCTGACAGAGGGGAGGACAGATC 9937  
Qy 433 CAGAGCTACGTGGCCATGTACAGCGGAACAGCGCTTGGAGAGGCGCTTGCAGCAGCTG 492  
Db 9936 CAGGAGGGGACGCGGGTGCAGGAGGGGACGCGGTGACAGAGGGGCGCGGCGCAGGAG 9877  
Qy 493 GCACGTGGGACACGCGCGGAGCCCTTGGAGG 526  
Db 9876 GGNACGGGGTGCAGGAGGGGACACAGTGCAGG 9843

## RESULT 8

US-09-954-456-1601  
Sequence 1601, Application US/09954456  
Patent No. US20020115057A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
FILE REFERENCE: 689290-76  
CURRENT APPLICATION NUMBER: US/09/954,456  
CURRENT FILING DATE: 2001-09-18  
PRIOR APPLICATION NUMBER: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,637  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,638  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720



;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: US/60/235,840  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: US/60/235,863  
;; PRIOR FILING DATE: 2000-09-27  
;; NUMBER OF SEQ ID NOS: 2276  
;; SOFTWARE: Patent in version 3.0  
;; SEQ ID NO 1601  
;; LENGTH: 14800  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-954-456-1601

Query Match 11.9%; Score 75.8; DB 10; Length 14800;  
Best Local Similarity 46.0%; Pred. No. 1.1e-08;  
Matches 294; Conservative 0; Mismatches 342; Indels 3; Gaps 1;  
  
Qy 1 GAGGTGAACGCCAGGCTAGCAGATCTGGTCTCAGGCACAGCTGAAGCACACGGG 60  
Db 5005 GAGGAGGAGAGCGGAGCTGGAGCGCTCAGGCGCAACGAGGCGCTACGGCTG 5064  
  
Qy 61 GCGAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTCGCACCAAG 120  
Db 5065 CGGCTGAGGCGGAGGAGTGGCGCAGCAGAGAGCTGGCGCAGCGCCAGCGCTGAGAG 5124  
  
Qy 121 GGCCTGGAGCTGGAGGCTGTGAGAAATGAGCTGAGCGCGCAAGAAAGACGAGCGGAGCTG 180  
Db 5125 CAGAAGGAGGAGCGGAGCGCGGCGCGCGCAAGGCGGAGGAGCGAGCGGCTC 5184  
  
Qy 181 CTGCGGAGAGGTGAACCTCTGAGCAGAGCTGAGGAGCTGGGGCCCGCCAGCGCC 240  
Db 5185 CGGAGCGGAGCTGGCTGAAACAGAGCTGGAGAGCAGCGGAGCTGGCGGAAGGCACC 5244  
  
Qy 241 CTGCGCGCGACATGGGGCGCCACCTTCCCGAGGAGCTCCCTGCGCTCGAGCGGAG 300  
Db 5245 GCGAGCAGCCCTGGCGCGGAGCAGAGTTGATCCGGCTGCGGCGCCGAGCAGCGAG 5304  
  
Qy 301 CTGAGCGCGCTGCGGCGCGAGCTGCGGAGGAGCGGCAAGGCCATGACAGATGTCTCG 360  
Db 5305 GGGAGCAGCAGCGCAGCTCTGAGGAGGAGCTGGCGCGCTGCGAGCGTGAGCGGCT 5364  
  
Qy 361 GGCTTCCAGCATGAGCGCTGTGTGAGAGGAGAGAGAGAGTGTATTCAGTACCAG 420  
Db 5365 GCAGCCAGCAGAAACCGAGGAGCTGGAGCCGAGCTGGCCCAAGGTG---CGGGCCGAG 5421  
  
Qy 421 AAACAGCTGCAGCAGAGCTAGCTGGCCATGTACAGCGGAACACGCGCTGGAGAGGCC 480  
Db 5422 ATGAGGTGCTGTGGCCAGCAAGCGGAGGCTGAGGAGGAGTCCGCTCCACAGCGAG 5481  
  
Qy 481 CTGAGCAGCTGGCAGCTGGGGACAGCGCCGGGAGCCCTTTGGAGGTTGACCTTGAAGGG 540  
Db 5482 AAGTCAAGCAGAGCTGGAGGCGGAGCGCGCGGCTTCCGCGAGCTGGCGGAGGAGGCC 5541  
  
Qy 541 CTGACATCCCTACGAGGACATATAGCCACTAGATCTGAGGGGCTGCTGGGAAGGC 600  
Db 5542 GCGCGCTGCTGCGCTTGGCGGAAGAGCCAAAGCGGAGCGGAGCTGGCGCGGAGAGAC 5601  
  
Qy 601 GAGTCTGGGACCTGGCACTGGAGGCGAGGCTCTCCCG 639  
Db 5602 GCGGCGCGGAGCGGCGGCGGAGGCGGAGGCGGCTTTCGCG 5640

RESULT 9  
US-09-854-133-157  
; Sequence 157, Application US/09854133  
; Publication No. US20020183499A1  
; GENERAL INFORMATION:  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Secrist, Heather  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

;; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
;; FILE REFERENCE: 210121.475C10  
;; CURRENT APPLICATION NUMBER: US/09/854,133  
;; CURRENT FILING DATE: 2001-05-11  
;; NUMBER OF SEQ ID NOS: 735  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 157  
;; LENGTH: 2313  
;; TYPE: DNA  
;; ORGANISM: Homo sapien  
US-09-854-133-157

Query Match 11.5%; Score 73.8; DB 9; Length 2313;  
Best Local Similarity 50.6%; Pred. No. 2.8e-08;  
Matches 208; Conservative 0; Mismatches 197; Indels 6; Gaps 1;  
  
Qy 77 TGGAGCTGAGGACCCAGGAGCTGGAGGGCGCCCTGCCACCAAGGGCTGGAGCTGGAGG 136  
Db 960 TGGCGATCCGGCTGCGAGCGCGGAGGAGAGCTCCCTACACAGGACCTGAAGCAGAAAC 1019  
  
Qy 137 TCTGTGAGAAATGAGCTGCAGCGCAAGAAAGAACGAGGCGGAGCTGCTCGGGAGAGGTGA 196  
Db 1020 GCGCGAGCAGCGGAGCAGCGGAGCGCGCGCGGCGG-----CCAGGAAGAGGAGC 1073  
  
Qy 197 ACTGTGAGAGCAGAGCTGCGAGAGCTGCGGGGCCAGGCGCCCTGCGCCCGACATGG 256  
Db 1074 TGCTGCGGCTGCAGCAGCTGCAGGAGGAGGAGCGGAGCTGCAGAGCTGGAGCTGC 1133  
  
Qy 257 GCGCGCCACCTTCCCGAGGAGCTCCCTGCTGCGAGCGGAGCTGGAGGGCTGCGGG 316  
Db 1134 TGCAGGAGGCGCAGCGCGAGCGCGCTGCTGCGAGGAGGAGGAGGAGGAGGAGGAGG 1193  
  
Qy 317 CCAGAGCTGCGGAGGAGCGGCAAGGCCATGACACAGATGTCTCGGGCTTCCAGCATGAGC 376  
Db 1194 GCGAGCAGCGGAGCTGCAGCAGCGCTCGAGGCGCAACTGCGCGAGCGGAGCAGGCC 1253  
  
Qy 377 GGCTGTGTGGAAGGAGGAGAGAGAGTGTATTCAGTACCAAGAAACAGCTGCAGCAGA 436  
Db 1254 GGGCTCATGCGAGCTGAGTGTAGTGAAGGAGGAGGAGGCTGCCCGGCGGCGCAGC 1313  
  
Qy 437 GCTACGTGGCCATGTACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 487  
Db 1314 GCATCAAGGAGCTGGAGGAGATGCAGCAGCGGTTGCGAGGAGGAGGAGGAGGAGGAGG 1364

RESULT 10  
US-09-738-973-157  
; Sequence 157, Application US/09738973  
; Patent No. US20020110563A1  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Fling, Steven P.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Algathe, Paul A.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Indirias, Carol Yoseph  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Elliott, Mark  
; APPLICANT: Mannion, Jane  
; APPLICANT: Kalos, Michael D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 210121.475C9  
; CURRENT APPLICATION NUMBER: US/09/738,973  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 587  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 157  
; LENGTH: 2313  
; TYPE: DNA  
; ORGANISM: Homo sapien

US-09-738-973-157

Query Match 11.5%; Score 73.8; DB 10; Length 2313;  
Best Local Similarity 50.6%; Pred. No. 2.8e-08;  
Matches 208; Conservative 0; Mismatches 197; Indels 6; Gaps 1;

QY 77 TGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCTGGAGCTGGAGG 136  
DB 960 TGGCGATCCGCTGACAGCCGAGGGGAGAGAGCTCCCTACACAGAGACCTGAAGCAGAAAC 1019

QY 137 TCTGTGAATAGCTGACGCGCAAGAAACGAGGCGGAGCTGCTCGGAGAGAGTGA 196  
DB 1020 GCGCGAGCAGCGGAGCAGCGGAGCGCGCGCGG-----CCAAGGAAGAGGAGC 1073

QY 197 ACCTGCTGGAGCAGGAGCTGACGAGCTGCGGGCCCGAGCGCCCTGCGCCCGCACATGG 256  
DB 1074 TGCTGCGGCTGCAGCAGCTGACGAGGAGAGAGGAGCGGAAGCTGCAGGAGCTGGAGCTGC 1133

QY 257 GCGCGCCACCTTCCCGCAGGACGTCCTGCTGCGCGGAGCTGGAGCGGCTGCGGG 316  
DB 1134 TGCAGGAGCGGAGCGGAGCGGAGCGGCTGCTGCGAGGAGGAGGAAACGCGCGCGCA 1193

QY 317 CCGAGCTCGGAGGAGCGGCAAGGCCATGACAGATGCTCTCGGGCTTCAGCATGAGC 376  
DB 1194 GCCAGCAGCGAGCTGCAGCAGCGCTGCGAGGGCCAACTGCGGAGGCGGAGCGGCC 1253

QY 377 GGCTCGTGTGAAGAGGAGAGAGAGAGAGTGTTCAGTACAGAAACAGCTGCAGCAGA 436  
DB 1254 GGGCTTCATCGAGCTGAGATGAGCTGGAAGGAGGAGGAGGCTGCCCGGAGCGGCGAGC 1313

QY 437 GCTAGCTGGCCATGTPACCAAGCGGACCGAGCGCTGGAGAGCGCCCTGCGAGC 487  
DB 1314 GCATCAAGAGCTGAGGAGATGACGAGCGGTTCAGGAGGCGCTTGCAC 1364

RESULT 11

US-09-764-868-12  
; Sequence 12, Application US/09764868  
; Patent No. US20020168711A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT232  
; CURRENT APPLICATION NUMBER: US/09/764,868  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1510  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 2314  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-868-12

Query Match 11.5%; Score 73.8; DB 9; Length 2314;  
Best Local Similarity 50.6%; Pred. No. 2.8e-08;  
Matches 208; Conservative 0; Mismatches 197; Indels 6; Gaps 1;

QY 77 TGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCTGGAGCTGGAGG 136  
DB 968 TGGCGATCCGCTGACGCGCGGAGGAGAGAGCTCCCTACACAGAGCTGAAGCAGAAAC 1027

QY 137 TCTGTGAATAGCTGACGCGCAGAGAGAGAGGCGGAGCTGCTCGGAGAGAGTGA 196  
DB 1028 GCGCGAGCAGCGGAGCAGCGGAGCGGAGCGCGCGCGG-----CCAAGGAAGAGGAGC 1081

QY 197 ACCTGCTGGAGCAGGAGCTGACGAGCTGCGGGCCCGAGCGCCCTGCGCCCGCACATGG 256  
DB 1082 TGCTCGGCTGCAGCAGCTGCAGGAGGAGAGGAGCGGAAGCTGCAGAGCTGGAGCTGC 1141

QY 257 GCGCGCCACCTTCCCGCAGGAGAGCTGCTGCGCTGCGCGGAGCTGCAGCGGCTGCGGG 316  
DB 1142 TGCAGGAGCGCAGCGGAGCGGAGCGGAGCGGCTGCTGCAGGAGGAGGAGGAAACGCGCGCGCA 1201

QY 317 CCGAGCTGCGGAGGAGCGGCAAGGCCATGACCAATGTCTCTCGGGCTTCAGCATGAGC 376  
DB 1202 GCCAGACCCCGAGAGCTGCAGCAGCGCTCGAGGGCCAACTGCGGAGCGGAGCAGGCC 1261

QY 377 GGCTCGTGTGCAAGGAGGAGAGAGAGAGTGTTCAGTACCAAGAAACAGCTGCAGCAGA 436  
DB 1262 GGGCTCCATGCAGGCTGAGATGAGCTGAAGGAGGAGGCTGCCCGGAGCGGCGAGC 1321

QY 437 GCTACGTGGCCATGTATCCAGCGGAAACAGCGCCCTGGAGAGAGGCCCTTCGAGC 487  
DB 1322 GCATCAAGGAGCTGGAGGAGATGCAGCAGCGGTTTCAGAGAGGCGCTTCAAC 1372

RESULT 12

US-09-925-299-215  
; Sequence 215, Application US/09925299  
; Publication No. US20030040617A9  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 215  
; LENGTH: 3323  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-299-215

Query Match 10.9%; Score 69.6; DB 9; Length 3323;  
Best Local Similarity 48.1%; Pred. No. 2.7e-07;  
Matches 232; Conservative 0; Mismatches 244; Indels 6; Gaps 1;

QY 65 AGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTGCGCACCAAGGGCC 124  
DB 1198 AGCTGGTGGCGCGGAGCAGGAGATCACGGCTGTGCGAGGACGCGATGCAGGCCAGCTACC 1257

QY 125 TGGAGCTGGAGGCTGTGAGAAATGAGCTGCAGCGCAAGAGAGAGCGGAGCTGCTGC 184  
DB 1258 GGGAGCACGTGAAGAGGCTGCAGCAGCTGCAGGGCAAGATCCCGAGCTCTTCAGGAGCAGC 1317

QY 185 GGGAGAAAGTGAACCTCTGAGCAGGAGCTGCAGGAGCTGCGGGCCCGAGCGCCCTTGG 244  
DB 1318 TGGAGATGGCCCCCAACACAGCAGCTGGCCCGCTGCAGCAGG-----AGNACTCCATCC 1371

QY 245 CCGCGCAGATGGGGCCGCCCACTTCCCGAGGAGCTCCCTGCGCTGCGAGCGGAGCTGG 304  
DB 1372 TCGGGGATGCTTGAACACGAGCCACGAGCCAGGCTGGAGAGCAAGCAGAGCTGG 1431

QY 305 AGCGCTGCGGGCGAGCTGCGGGAGGAGCGGAGGCCATGACCAATGCTCTCGGGCT 364  
DB 1432 CCAAGCTTCGCGAGGAGCTCAGCAAGGCTCAGCAAGAGCTGGTGGAGAGTCAAGGCTG 1491

QY 365 TCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGAGAGGTGATTTCAGTACCAAGAAC 424  
DB 1492 TCGGCGAAGATGACGACGAGCGGGAAGCTCTGAGCAAGGAGCTGCCTTCAGAGAGC 1551

QY 425 AGCTGCAGCAGAGCTACCTGGCCCATGTACCAGCGGAACCAAGCGCTTGGAGAGGCCCTTGC 484  
DB 1552 AGGTCTTCAGCTGACGCGCTCCACAGGAGAGTGAAGGAGGCGCTTGCAGAAAGCGCTGG 1611

QY 485 AGCAGCTGGCAGCTGGGGAGCAGCGCCCGGGAGCCCTTGGAGGTTGACTTGAAGGGGCTG 544  
DB 1612 ACAGAGTCAGCGGGAGCTGTGCCACACGAGCAGCAGCCAGCCAGCTCCGGGCGGATG 1671

QY 545 AC 546

Db 1672 CC 1673

## RESULT 13

US-09-925-299-215  
; Sequence 215, Application US/09925299  
; Patent No. US20020055627A1

## GENERAL INFORMATION:

; APPLICANT: Roben et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA102

; CURRENT APPLICATION NUMBER: US/09/925,299

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05883

; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1556

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 215

; LENGTH: 3323

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-925-299-215

Query Match 10.9%; Score 69.6; DB 10; Length 3323;

Best Local Similarity 48.1%; Pred. No. 2.7e-07;

Matches 232; Conservative 0; Mismatches 244; Indels 5; Gaps 1;

QY 65 AGCTGGAGGGCTTGGAGCTGAGGACCCAGGACCTGGAGGGGCGCTTGGCCACCAAGGGCC 124

Db 1198 AGCTGGTGGCCGGGAGCAGGAGATCACGGCTGTGCAGGCACGATGCGAGCCAGCTACC 1257

QY 125 TGGAGCTGAGGTCTGTGAGATGAGTGCAGCGCAAGAAACAGCGGGAGCTGTGC 184

Db 1258 GGGAGCAGCTGAAGAGGTGCAGCAGCTGCAGGGCAAGATCCGGACTCTTCAGGAGCAGC 1317

QY 185 GGGAGAAGGTGAACCTGTGGAGCAGGAGCTGCAGGAGCTGGGGCCCGGCGCTGG 244

Db 1318 TGGAGAATGGCCCAACACGAGCTGGCCCGCTGCAGCAGG-----AGAACTCCATCC 1371

QY 245 CCGGCGACATGGGGCGGCCACCTTCCCGAGGAGCTTCCCTGCCCTGCAGCGGGAGCTGG 304

Db 1372 TCGGGGATGCCTTGAACAGGCGCACGAGCTGAGAGCAAGCAGAAACGAGAGCTGG 1431

QY 305 AGCGCTCGGGCGGAGCTGGGGAGGAGCGGAGGCGCATCAGAGTGTCTCGGCT 364

Db 1432 CCAAGCTTCGGCAGGAGCTCAGCAAGGTTCAGAAAGAGCTGTGGAGAAGTCAGAGGCTG 1491

QY 365 TCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGGAGATTCAGTACCAGAAAC 424

Db 1492 TGGGCAAGATCAGCAGCAGCGGAAAGCTTGGAGCCAGGCAGCTGCCCTTCGAGAGC 1551

QY 425 AGCTGCACGAGCTACCTGGCCATGTACAGCGGAAACAGCGCTTGGAGAAGGCCCTTGC 484

Db 1552 AGGTCCTCAGCTGCAGGCGTCCACAGGAGAGTGCAGAGGCCCTTGCAGAAGCGCTTG 1611

QY 485 AGCAGCTGCGAGTGGGACAGCGCGGGAGCCCTTGGAGTTGACCTGGAAGGGGCTG 544

Db 1612 ACGAGTTCAGCCGGGAGCTGTGCCACACGAGAGCAGCCAGCCAGCCCTTCGGGGGGATG 1671

QY 545 AC 546

Db 1672 CC 1673

## RESULT 14

US-09-864-864-323  
; Sequence 323, Application US/09864864  
; Patent No. US20020102679A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Dillon, Davin C.

; APPLICANT: Secrist, Heather

; APPLICANT: Lodes, Michael J.

; APPLICANT: Algate, Paul A.

; APPLICANT: Fling, Steve P.

; APPLICANT: Mannion, Jane

; APPLICANT: Benson, Darin R.

; APPLICANT: Carter, Darrick

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER

; FILE REFERENCE: 210121.523

; CURRENT APPLICATION NUMBER: US/09/864,864

; CURRENT FILING DATE: 2001-05-23

; NUMBER OF SEQ ID NOS: 341

; SOFTWARE: Corixa Invention Disclosure Database

; SEQ ID NO 323

; LENGTH: 4372

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-864-864-323

Query Match 10.9%; Score 69.6; DB 10; Length 4372;

Best Local Similarity 48.1%; Pred. No. 2.7e-07;

Matches 232; Conservative 0; Mismatches 244; Indels 6; Gaps 1;

QY 65 AGCTGGAGGGCTTGGAGCTGAGGACCCAGGACCTGGAGGGGCGCTTGGCCACCAAGGGCC 124

Db 2977 AGCTGGTGGCCGGGAGCAGGAGATCACGGCTGTGCAGGCACGATGCGAGCCAGCTACC 3036

QY 125 TGGAGCTGGAGGTCTGTGAGAATGAGCTGCAGCGCAAGAAACAGAGCGGAGCTGTGC 184

Db 3037 GGGAGCAGCTGAAGGAGGTGCAGCAGCTGCAGGGCAAGATCCGGACTCTTCAGGAGCAGC 3096

QY 185 GGGAGAAGGTGAACCTGTGGAGCAGGAGCTGCAGGAGCTGGGGCCCGGCGCTGG 244

Db 3097 TGGAGAATGGCCCCCAACACGAGCTGGCCCGCTTCGACAGG-----AGAACTCCATCC 3150

QY 245 CCGGCGACATGGGGCGGCCACCTTCCCGAGGAGCTTCCCTGCCCTGCAGCGGGAGCTGG 304

Db 3151 TGGGAGTGCCTTGAACACAGCCACAGGAGCTGGAGAGCAGCAAGCAGAGAGCTGG 3210

QY 305 AGCGGCTCGGGCGGAGCTGGGGAGGAGCGGCAAGGCCATGACAGATGTCTCGGCT 364

Db 3211 CCAAGCTTCGGCAGGAGCTCAGCAAGGTTCACCAAGAGCTGTGAGAAGTCAGAGGCTG 3270

QY 365 TCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGGAGAGGTGATTCAGTACCAGAAAC 424

Db 3271 TGGGCAAGATGAGCAGCAGCGGAAAGCTTGGAAAGCCAGGCAGCTGCCCTTCGAGAAAGC 3330

QY 425 AGCTGCAGCAGAGCTACGTGGCCATGTACACGCGGAACACAGCGCTTGGAGAAGGCCCTTGC 484

Db 3331 AGGTCCTGCAGCTCAGCGCTCCACAGGGAGAGTGGAGAGGCCCTTGCAGAAGGCCCTTGC 3390

QY 485 AGCAGCTGGCAGCTGGGACACAGCGCCGGGAGCCCTTGGAGTTGACCTGGAAGGGGCTG 544

Db 3391 ACGAGGTCAGCCGGAGCTGTGCCACACGAGCAGCAGCCAGCCAGCTTCCGGGGGGATG 3450

QY 545 AC 546

Db 3451 CC 3452

## RESULT 15

US-09-864-864-337  
; Sequence 337, Application US/09864864  
; Patent No. US20020102679A1

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Dillon, Davin C.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 889.672 Seconds  
(without alignments)

11632.279 Million cell updates/sec

Title: US-09-513-888C-1\_COPY\_4912\_5550

Perfect score: 639

Sequence: 1 gaggtgaacccaaggtag.....tggaggcagggtctctcccg 639

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	175.8	27.5	247	9 AL135213	AL135213 DKFZp762E
2	171.2	26.8	326	10 AW457817	AW457817 UI-M-BH3-
3	159.4	24.9	795	13 BI335611	BI335611 EST00009
4	131.8	20.6	562	12 BF079806	BF079806 230497 MA
5	128.8	20.2	824	13 BI655319	BI655319 603283348
6	125.2	19.6	1053	14 BQ431916	BQ431916 AGENCOURT

7	121.2	19.0	918	12	BG385903	BG385903
8	120.4	18.8	482	13	BM152503	BM152503 TCBAPIE91
9	120	18.8	453	14	BM855056	BM855056 K-EST0137
10	120	18.8	901	12	BF317160	BF317160 601901286
11	120	18.8	926	12	BG035858	BG035858 602325955
12	120	18.8	957	12	BG170653	BG170653 602323523
13	118.6	18.6	923	12	BG750395	BG750395 602709229
14	118.4	18.5	469	13	BM193264	BM193264 TCBAPIE61
15	118.2	18.5	450	14	BQ694854	BQ694854 1000997 H
16	117.2	18.3	705	13	BI104307	BI104307 602889876
17	117	18.3	357	14	BM854246	BM854246 K-EST0136
18	116.2	18.2	806	13	BM005722	BM005722 603613327
19	115.2	18.0	914	13	BI766138	BI766138 603052944
20	114.8	18.0	977	12	BF032717	BF032717 601453408
21	114.4	17.9	799	13	BI146842	BI146842 602911741
22	114.2	17.9	898	14	BQ646816	BQ646816 AGENCOURT
23	114.2	17.9	940	12	BG830547	BG830547 602767136
24	113	17.7	718	17	AZ972907	AZ972907 2M0246M24
25	112.4	17.6	704	13	BM042462	BM042462 603616294
26	112.2	17.6	342	13	BM106435	BM106435 510165 MA
27	111.4	17.4	883	14	BQ946535	BQ946535 AGENCOURT
28	110.8	17.3	858	13	BI250802	BI250802 602993419
29	110.4	17.3	287	14	BM818585	BM818585 K-EST0085
30	107.4	16.8	839	13	BI752956	BI752956 603025713
31	107	16.7	811	17	CNS038X0	AL233181 Tetradon
32	106.6	16.7	973	9	AL551801	AL551801
33	106.4	16.7	432	12	BF313433	BF313433 601900020
34	106.4	16.7	1106	14	BQ648477	BQ648477 AGENCOURT
35	105.6	16.5	587	12	BF309857	BF309857 601892321
36	105.2	16.5	650	17	AZ337339	AZ337339 1M0068803
37	102.4	16.0	605	17	AZ640388	AZ640388 1M0502E21
38	101	15.8	751	10	BE312985	BE312985 601150246
39	101	15.8	932	14	BQ879505	BQ879505 AGENCOURT
40	100.2	15.7	395	13	BM146503	BM146503 TCAAP1E49
41	100	15.6	564	13	BJ098506	BJ098506 BJ098506
42	94.8	14.8	486	9	AI605748	AI605748 ma66d08.Y
43	94.8	14.8	811	12	BG866479	BG866479 602785427
44	94.2	14.7	713	17	AZ346186	AZ346186 1M0081G06
45	93.8	14.7	535	10	AW420993	AW420993 fj90d02.x

#### ALIGNMENTS

RESULT 1  
AL135213  
LOCUS AL135213 247 bp mRNA linear EST 25-FEB-2000  
DEFINITION DKFZp762E2313\_r1 762 (synonym: hmel2) Homo sapiens cDNA clone  
ACCESSION AL135213  
VERSION AL135213.1 GI:6603400  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 247)  
AUTHORS Ansong, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.  
TITLE EST (Ansong, W., Wirkner, U., Mewes, H.W., Weil, B. and Wiemann, S.)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Ansong W  
MIPS Am Klopferspitz 18a D-82152 Martinsried, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No s1 sequence available.  
This clone (DKFZp762E2313) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.



Meleagris.  
1 (bases 1 to 795)  
Harry D.E., Zaitlin, D., Marini, P.J. and Reed, K.M.  
A first-generation map of the turkey genome  
Unpublished (2001)  
Contact: Reed KM  
Department of Veterinary Pathobiology,  
University of Minnesota  
295 AS/VM, 1988 Fitch Ave, St. Paul, MN 55108, USA  
Tel: 612 624 1287  
Fax: 612 625 0204  
Email: reedx054@cc.umn.edu  
Seq primer: TCGAAGTTCCTCCCTCCTCACTAAAGG  
POLYA-No.

FEATURES  
source  
1..795  
/organism="Meleagris gallopavo"  
/db\_xref="taxon:9103"  
/clone="Nte083"  
/clone\_lib="Turkey Lambda ZAP Library"  
/sex="unknown"  
/tissue\_type="whole embryo"  
/dev\_stage="24-day"  
/note="Vector: Lambda ZAP Library ZAP; RNA was isolated from a 24-day turkey embryo using a commercially available kit (Stratagene). cDNA library was constructed using a Lambda ZAP (SK-, Stratagene)."

BASE COUNT  
188 a 232 c 214 g 161 t

Query Match 24.9%; Score 159.4; DB 13; Length 795;  
Best Local Similarity 77.5%; Pred. NO. 5.7e-19;  
Matches 193; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 334 CGGAAGCCATGACCATGCTCCGGCTTCAGCATGAGCGCTCGTGTGAAGGAG 393  
Db 1 CGGACACACGACGATGATGCTCCAGCTTCAGCAGGAGCGGACATGAGGGAG 60

Qy 394 GAGAAGGAGAGGTGATTTCAGTACAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTAC 453  
Db 61 GAAAAGAGAGGTGATCCATACAGAAAGCAGTTGCAGCAGAGCTACCTCATGTAC 120

Qy 454 CAGCGAACCAGCGCTGGGAAGGCCCTGCAGCAGTGGCAGCGTGGGACAGCGCCGGG 513  
Db 121 AAGAGAACCAAGAACCTGGAGAAGATGCTGCAGCAGCTCGCGGGGGGAGGTGCAAG 180

Qy 514 GAGCCCTGGAGGTGACTGGAAGGGGCTGCATCCCTACGAGGACATCATAGCCACT 573  
Db 181 GAGCCCAATTGAGCTGGACATCCCGGCGCTGACGTGCTCCCTACGAGGACATCATAGCCACT 240

Qy 574 GAGATCTCA 582  
Db 241 GAGATCTCA 249

RESULT 4  
BF079806 562 bp mRNA linear EST 18-OCT-2000  
LOCUS BF079806  
DEFINITION 230497 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION BF079806  
VERSION BF079806.1 GI:10873636  
KEYWORDS EST.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
1 (bases 1 to 562)  
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,  
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,  
and Keele, J.W.  
Design and use of two pooled tissue normalized cDNA libraries for  
EST discovery in swine  
Unpublished (2000)

CONTACT: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@mail.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -minmatch 12 options.  
PCR Primers  
FORWARD: AGGAAACAGCTATGACCAT  
BACKWARD: GTTTCCCGAGTCAGCAGC  
Plate: 47 row: K column: 8  
Seq primer: ATTAGGTGACACTATAG.  
FEATURES  
source  
1..562  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_lib="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH108"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from testis, ovary, and  
endometrium, pituitary, and placenta."  
BASE COUNT 106 a 165 c 226 g 65 t

ORIGIN  
Query Match 20.6%; Score 131.8; DB 12; Length 562;  
Best Local Similarity 54.9%; Pred. No. 4.5e-14;  
Matches 308; Conservative 0; Mismatches 217; Indels 36; Gaps 1;

Qy 20 GCGAGATCTCTGGGTCTCAAGGCACAGCTGAAGGACACGCGGGGCAAGCTGGAGGCGCTGG 79  
Db 1 GCGAGCTGGTAGCCTCGAGTGGCGCTACGGGAGGCCCGGCGAGCGCTGCGGTGAGTG 60

Qy 80 AGCTGAGACCCAGGACCTGGAGGCGCCTTGCGCACCAAGGGCGCTGAGCTGAGGTCT 139  
Db 61 AGGCGCGGGCGGGGCTCCAGGAGGCCGCCCAACACGCGAGCTGAGGTGGAGGCGCT 120

Qy 140 GTGAGATGAGTGCAGCGCAAGAAAGACGAGGCGGAGCTGCTCGGGGAGAGGTGAACC 199  
Db 121 GTTCCCAAGAGCTGCAGCGGCATGCCAGAGGCTGAGCGGCTCGAGAGAAAGCGGGAC 180

Qy 200 TGCTGGAGCAGGAGCTGCAGGAGCTGCGGGC----- 230  
Db 181 AGTTGGACACCGAGGCGGTGACTCGGGAAACCCCTTGTGCCACTTGCCACTGCCGACC 240

Qy 231 -----CGAGGCGCCCTGCGCCGACATGGGGCGCCCACTTCCCGAGGACGTCC 283  
Db 241 CATTCCTCTGGCAGAGGCGATGAAGCAAGGCACAGCGGCGAGCTGCTGGGGTTGGGG 300

Qy 284 CTGCGCTGCAGCGGAGCTGGAGCGGCTGCGGGCGGAGCTGCGGAGAGCGGCAAGGCC 343  
Db 301 CGAGCCTGCGGGCGCCAGGTGGAGCGTCTCGGGGAGATCTGACGCGGAGCGGGCGGTG 360

Qy 344 ATGACCAGATGTCCTCGGGCTTCAGCATGAGCGGCTCGTGTGAAGAGAGAGAGAGA 403  
Db 361 GCGAGGAGCAGCGGACAGCTTTGAGGGGAGCGGCTGGCTGCGAGCGCGAGAGAGC 420

Qy 404 AGGTGATTTCAGTACCAAGAAACAGCTGAGCAGAGCTACGTGGGCCATGTACCGGGAGC 463  
Db 421 AGGTGATCCGCTACCAAGAGCAGCTGCAGCACAACATACATCCAGATGTACCGAGCAACA 480

Qy 464 AGCGCTGGAGAGGCGCTGCAGCAGCTGGGACATGCGGGACAGCGCGGGAGCGCTTGG 523  
Db 481 GGCAGCTGGAGAGGAGCTACAGCAGCTCAGCTGGAGCTGGAGGCGCGGAACTCGCCG 540

Qy 524 AGGTGACCTGGAGAGGCGGTG 544  
Db 541 ATCTGGCGCTGGCGAGGCCAG 561

RESULT 5

```

BI655319          824 bp      mRNA      linear      EST 12-SEP-2001
LOCUS             603283348F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5327843 5',
DEFINITION        mRNA sequence.
ACCESSION         BI655319
VERSION           BI655319.1 GI:15569555
KEYWORDS          EST.
SOURCE            house mouse.
ORGANISM          Mus musculus.
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS           NIH-MGC http://mgi.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
                  Ph.D.
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM11831 row: o column: 12
                  High quality sequence stop: 823.
FEATURES          source
                  Location/Qualifiers
                    1..824
                    /organism="Mus musculus"
                    /strain="NMRI"
                    /db_xref="taxon:10090"
                    /clone="IMAGE:5327843"
                    /clone_lib="NCI_CGAP_Mam4"
                    /tissue_type="tumor, gross tissue"
                    /dev_stage="5 months"
                    /lab_host="DH10B"
                    /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
                    Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
                    Library constructed by Life Technologies. Investigators
                    providing samples: Lothar Hennighausen/Priscilla Furth,
                    NIH Reference for transgenic model: Li et al., Cell Growth
                    and Differentiation 7, 3-11 (1996)."
```

```

BASE COUNT       156 a 251 c 295 g 122 t
ORIGIN
Query Match      20.2%; Score 128.8; DB 13; Length 824;
Best Local Similarity 55.3%; Pred. No. 1.6e-13;
Matches 321; Conservative 0; Mismatches 217; Indels 42; Gaps 2;
QY 58 CGGGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGGCGACC 117
DB 42 CGGGCCACCTTGGCGGGTCACTGAGGGCGGTGCCAGGGCCTTCAGAGGGCAGCCGAGCT 101
QY 118 AAGGCGCTGGAGCTGGAGCTCTGAGAGATGAGCTGACGGCAAGAACGAGCGGAG 177
DB 102 CGGAGCAGGAGCTGGAGGCTCTGCGAGGAGCTTCAGCGGTACCGGCAGGAGGCTGAG 161
QY 178 CTGCTCGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGGGC----- 230
DB 162 CGGCTCCGGGAGAGCGGGGCAATTGGATGCGGAAGCATCGGACTCCGGGATCCCT 221
QY 231 -----CGAGCGCGCCTGGCCCGCGACATGGGGCG 261
DB 222 GTACCCCTGCCACACCGACCCGTTCTCTTGGCAGAGAGTGTATGAGGCCAAGGTGCAG 281
QY 262 CCACCTTCCCGAGAGCTCCCTGCCTGCGAGCGGGAGCTGGAGCGCTGCGGGCCGAG 321
DB 282 CGGGCAGCGGCTGGGGCAGGGGGTAGCTGCGGGCTCAGGTGGAGAGCTGCGCCAGAA 341
QY 322 CTGCGGGAGGAGCGGCAAGGCCATGACACAGATGTCTCTGGGCTTCCAGATGAGCGGCTC 381
DB 342 CTGCAGCGGGAGCAGCGGCGAGGGGACGAGCAACGGGACAGCTTTGAGGGGAGCGGCTG 401
```

```

QY 382 GTGTGAAGCAGGAGGAGAGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAGCTAC 441
DB 402 GCCTGGCAGCAGAGAGAGGACAGGTGATCCGCTACCAGAGCAGCTGCAGCAACTAT 461
QY 442 GTGGCCATGTACCAGCGGAACACAGCGCTGGAGAAAGCCCTGCAGCAGCTGGCAGCTGGG 501
DB 462 ATCCAGATGTACCGGCTTACCGCGAGCTGGAGCAGGAACCTGCAGCAACTGAGCTGGAG 521
QY 502 GACAGCCCGGGAGAGCCCTTGGAGTTGACTGTGAAGG-----GCTGACATCCCCCTAC 555
DB 522 CTGAGGCGCGCAGCTTGGGATTTGGGCTGGCGAGTCGGCCCTTCATCTGCTGTG 581
QY 556 GAGGACATCATAGCCACTGAGATCTGAGGGGCTGCCTGGG 595
DB 582 GAGGAGATCACTGCCACTGAGATCTAGGGCTCTGCAAGGG 621

RESULT 6
LOCUS             B0431916          1053 bp      mRNA      linear      EST 24-MAY-2002
DEFINITION        AGENCOURT 7898255 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6154088
                  5', mRNA sequence.
ACCESSION         B0431916
VERSION           B0431916.1 GI:21170992
KEYWORDS          EST.
SOURCE            human.
ORGANISM          Homo sapiens
REFERENCE         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS           NIH-MGC http://mgi.nci.nih.gov/.
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Life Technologies, Inc.
                  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Agencourt Bioscience Corporation
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM13494 row: j column: 09
                  High quality sequence stop: 378.
FEATURES          source
                  Location/Qualifiers
                    1..1053
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:6154088"
                    /clone_lib="NIH_MGC_67"
                    /tissue_type="retinoblastoma"
                    /lab_host="DH10B (phage-resistant)"
                    /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
                    Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                    Average insert size 1.75 kb. Library constructed by Life
                    Technologies."
BASE COUNT       179 a 349 c 377 g 148 t
ORIGIN
Query Match      19.6%; Score 125.2; DB 14; Length 1053;
Best Local Similarity 70.9%; Pred. No. 7.5e-13;
Matches 166; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 286 GCCTTGAGCGGGAGCTGGAGCGGCTCGGCGCCAGCTGCGGGAGGAGCGCAAGGCCAT 345
DB 187 GCCTTGGCGGGAGGTGGGGCGCTGCAGCCGAGCTGCGGCTGAGCGCGGCGCGG 246
QY 346 GACCAGATGTCTCGGGCTCCACATCAGCGGCTCGTGTGGAAGGAGGAAGAGAGAG 405
DB 247 GAGCGCAGGGTGCAGCTTCGCCGAGAGCGCCGCTGTGGTGGAGAGAGAGAGAG 306
QY 406 GTGATTCACTACAGAAACAGCTCAGCAGAGCTACGTGGCCATGTATTACCGGAACCA 465
```



Query Match	18.8%	Score 120.4;	DB 13;	Length 482;
Best Local Similarity	58.2%;	Pred. NO. 4.7e-12;		
Matches 233:	Conservative	0;	Mismatches 161;	Indels 6;
				Gaps 1;

```

Qy 230 CCCAGGCGCCCTGGCCCGGACATGGGGCGGCCACCTTCCCGAGGACGTCCCTGCC 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 CCCTGGCAGAGTGTATGAGCCAAAGTGCAGCGGCGAGCAGCCGGGTGGGGCAGCT 422

Qy 290 TGCAGCGGGAGCTCGAGCGGCTGCGGGCCGAGCTGCGGGAGGAGCGCAAGGCCATGACC 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 TGCAGCGGGAGCTCGAGCGGATTCGGGTGGAGCTGCAGCGGAGCGCGGGGGTGGAGG 362

Qy 350 AGATGTCTCTGGGCTTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGAAGAGTGA 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 AGCAGCGGGACAGCTTTGAGGGGAGCGGCTGGCCCTGCGCAGGACAGAGAGAGTGA 302

Qy 410 TTCAGTACCAAGAACAGCTGCAGCAGAGTACGTGGCCATGTACACGCGGAACAGCGCC 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TCCCTACCAAGAACAGCTGCAGCAGACATATACATCCAGATGTACCGGCGCAACCGGAGC 242

Qy 470 TGAAGAAGGCCCTGCAGCAGCTGCAGCTGGGAGCAGCGCGGGGAGCCCTTGGAGGTG 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TAGAGCAGGAGCTGCAGCAGCTCAGCCTGGAGCTGGAGGCCCGGAGCTCGCTGACCTGG 182

Qy 530 ACCTGG-----AAGGGCTGACATCCCTACGAGGACATATAGCCACTGAGATCTGAG 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GCCTGGCGGAGCAGGCGCCCTGCTGCTGCTGGAGGAGATCACTGCTACTGAGATATAGG 122

Qy 584 GGGCTGCTGGGAAGGCGAGTCTGGGACCTGGCACTGGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCCCTCAGCAACAGCTCTGTAGGAGCTCTGCCAGAGGG 82

```

```

RESULT 9
BM855056
LOCUS K-EST0137829 S21SNU520 Homo sapiens cDNA clone S21SNU520-62-E11 5',
DEFINITION mRNA sequence.
ACCESSION BM855056
VERSION BM855056.1 GI:19211455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 453)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 62 row: E column: 11
High quality sequence stop: 453.
Location/Qualifiers
1. 453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S21SNU520-62-E11"
/clone_lib="S21SNU520"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/lab_host="SNU-520"
/lab_host="Top10P"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR

```

```

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 62 row: E column: 11
High quality sequence stop: 453.
Location/Qualifiers
1. 453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S21SNU520-62-E11"
/clone_lib="S21SNU520"
/sex="F"
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/lab_host="SNU-520"
/lab_host="Top10P"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR

```

```

FEATURES
source

```

I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transposition of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

```

BASE COUNT 92 a 129 c 173 g 59 t
ORIGIN

```

```

Query Match 18.8%; Score 120; DB 14; Length 453;
Best Local Similarity 59.6%; Pred. No. 5.5e-12;
Matches 223; Conservative 0; Mismatches 145; Indels 6; Gaps 1;

Qy 256 GGGCGGCCACCTTCCCGAGGAGCTCCCTGCCCTGCAGCGGAGCTGGAGCGGCTGCGG 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 GTGAGCGGCGCAGCAGCCGGGTGGGGCAGCTTGGGGCCCGAGTGGAGCGATTGCGG 63

Qy 316 GCCAGCTGCGGAGGAGCGCAAGCCATGACCATGTCTTGGGCTTTCCAGCATGAG 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GTGAGCTGCAGCGGAGCGCGGGGTGAGGAGCAGCGGAGCAGCTTTAGGGGGAG 123

Qy 376 CGGCTCGTGTGAAGGAGGAGGAGAGTGTATTCAGTACCAGAAACAGCTGCAGCAG 435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 CGGCTGGCCTGGCAGGAGGAGGAGGAGTATCCGCTACCAGAAAGCAGCTGCAGC 183

Qy 436 AGCTACGTGGCCATGTACACGCGGAACACGCGCTGGAGAAGGCCCTGCAGCAGCTGGCA 495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 AACTACATCCAGATGTACCGCGCCACCGCAGCTAGAGCAGAGCTGCAGCAGCTCAGC 243

Qy 496 CGTGGGAGACGCGCGGGAGCCCTTGGAGGTTGACCTGG-----AAGGGGCTGACATC 549
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 CTGGAGCTGGAGGCCCGGGAGCTCGCTGACCTGGGCTGGCGGAGCGGCGGCTGCATC 303

Qy 550 CCTACGAGGACATCATAGCAGCTAGATCTGAGGGGCTGCCCTGGGAGGCGAGCTCTGGG 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 TGCTTGAGGAGATCACTGTCTAGATCTAGGGCCCTCAGCAACCAAGCTCTCTAGGGA 363

Qy 610 GACCTGGCACTGGG 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 GCTCTGCCAGAGGG 377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
BF317160
LOCUS BF317160
DEFINITION 601901286F1 NIH MGC_19 Homo sapiens cDNA clone IMAGE:4130469 5',
mRNA sequence.
ACCESSION BF317160
VERSION BF317160
KEYWORDS EST.
SOURCE BF317160.1 GI:11265566
human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 901)
NTH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM1026 row: 1 column: 22
High quality sequence start: 2

```

```
High quality sequence stop: 589.
Location/Qualifiers
1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4130469"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab host="DHI0B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into ECORI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 199 a 247 c 311 g 144 t
ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 901;
Best Local Similarity 100.0%; Pred. No. 6.2e-12;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 579
DB 3 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 62
QY 580 TGAGGGCTGCTGGGAGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 639
DB 63 TGAGGGCTGCTGGGAGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 122

RESULT 11
BG035858
LOCUS 602325955F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4414260 5',
DEFINITION mRNA sequence.
ACCESSION BG035858
VERSION BG035858.1 GI:12430413
KEYWORDS EST...
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 926)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10141 row: m column: 13
High quality sequence stop: 735.
Location/Qualifiers
1..926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4414260"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab host="DHI0B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 199 a 276 c 326 g 161 t
ORIGIN

High quality sequence stop: 589.
Location/Qualifiers
1..901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4130469"
/clone_lib="NIH_MGC_19"
/tissue_type="neuroblastoma"
/lab host="DHI0B (phage-resistant)"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into ECORI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 199 a 247 c 311 g 144 t
ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 901;
Best Local Similarity 100.0%; Pred. No. 6.2e-12;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 579
DB 3 TTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATC 62
QY 580 TGAGGGCTGCTGGGAGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 639
DB 63 TGAGGGCTGCTGGGAGGGAGTCTGGGACCTGGCACTGGGAGGAGGGCTCTCCG 122

RESULT 12
BG170653
LOCUS 602325232F1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426924 5',
DEFINITION mRNA sequence.
ACCESSION BG170653
VERSION BG170653.1 GI:12677356
KEYWORDS EST...
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 957)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLAM10174 row: m column: 05
High quality sequence stop: 681.
Location/Qualifiers
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4426924"
/clone_lib="NIH_MGC_89"
/tissue_type="hypermephroma, cell line"
/lab host="DHI0B (phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 194 a 276 c 326 g 161 t
ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 926;
Best Local Similarity 61.7%; Pred. No. 6.2e-12;
Matches 211; Conservative 0; Mismatches 125; Indels 6; Gaps 1;

QY 288 CCTCAGCGGAGCTGGAGCGGCTGCGGGCCGAGCTGCGGAGGAGCGCGGAGGCGCATGA 347
DB 5 CTTGGGGCCCGAGCTGGAGCGGATTGCGGGTGAGCTGCAGCGGAGCGCGCGGGTGA 64
QY 348 CCAGATGTCCTCGGGCTTCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGAGGT 407
DB 65 GGAGCAGCGGCGACAGCTTTGAGGGGAGCGGCTGCGCTGGCAGGCGAGAGGAGCAGGT 124
QY 408 GATTGAGTACAGAAACAGCTGCAGCAGAGCTAGCTGGCCATGTACCAGCGGAACCGCG 467
DB 125 GATCCGCTACCAAGAGCAGCTGCAGCAACTATCATCCAGATGTACCGCGCAACCGCA 184
QY 468 CCTGGAAGGCGCTGCAGCAGCTGGCACGTGGGAGACAGCGCGGGAGCGCTTGGAGGT 527
DB 185 GCTAGAGCAGAGCTGCAGCAGCTCAGCTGGAGCTGGAGCGCGGAGCTCGCTGACCT 244
QY 528 TGACCTGG-----AAGGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTG 581
DB 245 GGGCTGTGGCCGAGCAGGCGGCTGCTGCTGGAGGAGATCACTGCTACTGAGATCTA 304
QY 582 AGGGGCTGCTGGGAGGCGAGTCTGGGGACCTGGCACTGGG 623
DB 305 GGGCCCTCAGCAACAGCTCTGTAGGGAGCTCTGCCAGAGGG 346

FEATURES
source
1..957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4426924"
/clone_lib="NIH_MGC_89"
/tissue_type="hypermephroma, cell line"
/lab host="DHI0B (phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

BASE COUNT 194 a 276 c 326 g 161 t
ORIGIN
```

## ORIGIN

Query Match 18.8%; Score 120; DB 12; Length 957;  
 Best Local Similarity 61.7%; Pred. No. 6.2e-12;  
 Matches 211; Conservative 0; Mismatches 125; Indels 6; Gaps 1;  
 QY 288 CTTGAGCGGAGCTGGAGCGGCTCGCGGCGAGCTGCGGAGGAGCGCAAGGCCATGA 347  
 Db 8 CTTGCGGCGCCAGCTGGAGCGATTCGCGGTGGAGCTGACGCGGAGCGCGCGGGGTGA 67  
 QY 348 CCAGATGCTCTGGGCTTCCAGCATGACGCGCTCGTGTGGAAGGAGGAGGAAGGT 407  
 Db 68 GGAGCAGCGGACAGCTTTGAGGGGAGCGCGCTGCGCTGGCAGGACAGAGAAGGACAGT 127  
 QY 408 GATTGATACAGAAACAGCTGCAGCAGAGCTAGCTGCGCATGTACACAGCGGAACACGCG 467  
 Db 128 GATCGCTTACAGAAAGCAGCTGCAGCAACACTACATCCAGATGTATCCGCGCAACCGGCA 187  
 QY 468 CTTGAGAAAGCGCTGCAGCAGCTGGCACGTGGGGACAGCGCGCGGGAGCGCTTGGAGGT 527  
 Db 188 GCTAGAGCAGAGCTGCAGCAGCTCAGCTGGAGCTGGAGCGCGGAGCTGCTGACCT 247  
 QY 528 TGACCTGG-----AAGGGGTGACATCCCTACGAGACATCATAGCCATGAGATCTG 581  
 Db 248 GGGCTTGGCGGAGCAGCGGCGCTGCTGCTGCTGGAGGAGATCACTGCTACTGAGATCTA 307  
 QY 582 AGGGGCTCCCTGGGAGCGAGCTGCGGAGCTGCGACTGGG 623  
 Db 308 GGGCGCTCAGCAACAGCTCTGTAGGAGCTCTGCCAGAGG 349

## RESULT 13

LOCUS BG750395 923 bp mRNA linear EST 15-MAY-2001  
 DEFINITION 602709229F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4845799 5',  
 mRNA sequence.

## ACCESSION

VERSION BG750395.1 GI:14061048

## KEYWORDS

EST.

## SOURCE

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 923)

NIH-MGC <http://mgi.nci.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

CNA Library Preparation: Ling Hong/Rubin Laboratory

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

plate: LCM1685 row: j column: 08

High quality sequence stop: 895.

Location/Qualifiers

1..923

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4845799"

/clone\_lib="NIH\_MGC\_43"

/tissue\_type="normal pigmented retinal epithelium"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library. |"

BASE COUNT 180 a 246 c 368 g 129 t

Query Match 18.6%; Score 118.6; DB 12; Length 923;  
 Best Local Similarity 52.8%; Pred. No. 1.1e-11;  
 Matches 307; Conservative 0; Mismatches 239; Indels 35; Gaps 1;

QY 1 GAGTGTAACGCCAAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAGCACACGCGG 60  
 Db 269 GAGCTGGTGCAGAAAGGCGCAGCTGGTGTCTCTCGGGTGGCGCTGCGGAGGCGCGT 328  
 QY 61 GGCAAGCTGGAGGCGCTGGAGCTCAGGACCCAGACCTGGAGGGCGCCCTCGCACCAAG 120  
 Db 329 GCTACGCTGCGGCTCAGTGAGGCGCTGCGGCGGTCTACAGGAGGCCGCCGAGCTCGG 388  
 QY 121 GGCCTGGAGCTGGAGGCTGTGTAGAATGAGCTGCAGCGCAAGAAAGAACAGCGGAGCTG 180  
 Db 389 GAGCTGGAGCTGGAAAGCTGTTCACAGAGCTGCAGCGACACCGCCAGGAAGCTGAGCAG 448  
 QY 181 CTGCGGAGAGAGGTGAACCTGCTGGAGCAGAGCTGCAGGAGCTGCGGGCCCA----- 233  
 Db 449 CTGCGGAGAGAAAGCTGGGCACTTGGATGCTGAGCGCGGCACTCCGGGAGCCCTCTGTGC 508  
 QY 234 -----GGCGGCCCTGGCGCGACATGGGCGCGCCCA 265  
 Db 509 CACCTGCCACCGTTGACCCATTCTCTGGCAGAGAGTGATGAGGCCAAAGTCAGCGGG 568  
 QY 266 CTTTCCCGAGGACGTCCTGCTGCAGCGGGAGCTGGAGCGCTGCGGGCCGAGCTGTC 325  
 Db 569 CAGCAGCGGGGTTGGGGGAGCTTGGCGGCGCCAGGTGGAGCGATTGGCGGTGGAGCTGC 628  
 QY 326 GGGAGGAGCGGCAAGGCCATGACCAGATGCTCTCGGGCTCCAGCATGAGCGGCTCGTGT 385  
 Db 629 AGCGGAGCGCGCGCGGGGTGAGGAGCAGCGCGGACAGCTTTGAGGGGAGCGGCTGG 688  
 QY 386 GGAAGGAGGAGAGGAGAGGTGATTACGTACCAAGAACAGCTGCAGCAGAGCTACGTGG 445  
 Db 689 GGAGGAGCAGAGAGAGAGAGGTGATCGCTACCAAGAGCAGCTGCAGCAGCACTACATCC 748  
 QY 446 CCATGTACCGGAAACCGAGCGCTTGGAGAGAGGCGCTTCAGCAGCTGCACAGCTGGGACA 505  
 Db 749 AGATGTACCGGCGCAACCGGCGCAGTAGAGCAGGAGCTGAGCAGCTCAGCTGGAGCTTG 808  
 QY 506 GCGCGGGGAGCGCTTGGAGGTTACCTGGAGGGGCTGAC 546  
 Db 809 GAGGCGCGGAGCTCGTTGACCTGGGCTTGGCGGAGCAGGC 849

## RESULT 14

BM193264/c

LOCUS BM193264

DEFINITION TCBAP1E6182 Pediatric pre-B cell acute lymphoblastic leukemia

Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP6182, mRNA

sequence.

ACCESSION BM193264

VERSION BM193264.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 469)

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,

Gunarathne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project (2001)

Unpublished (2001)

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

## FEATURES

source

FEATURES  
source  
Email: clones@txccc.org  
Seq primer: M13 primer:  
Location/Qualifiers  
1. .469  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clones="TCBAP182"  
/clone\_lib="Pediatric pre-B cell acute lymphoblastic  
leukemia Baylor-HGSC project=TCBA"  
/sex="male"  
/tissue\_type="leukopheresis"  
/cell\_type="pre-B cell"  
/dev\_stage="pediatric 2 years"  
/lab\_host="DH10B"  
/note=vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;  
First strand cDNA was primed with an anchored  
XhoI-oligo(dT) primer [5'-GGAGGACTGAGCGCGCAGGAGG(T)VN  
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second strand  
was primed with a BamHI-dC primer  
[5'-AGAGGCTCGATCGCGCGCCGAATATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI  
and directionally cloned into the BamHI and SalI sites of  
lambda pSB vector. Library went through one round of  
normalization. Library was constructed by Wei Yu at RIKEN  
of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,  
Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,  
Schneider C, Hayashizaki Y, High efficiency selection of  
full-length cDNA by improved biotinylated cap trapper.,  
DNA Res 4: 1, 61-6, Feb 28, 1997)"  
a 175 c 132 g 99 t

BASE COUNT  
ORIGIN

Query Match 18.5%; Score 118.4; DB 13; Length 469;  
Best Local Similarity 59.4%; Pred. No. 1.1e-11;  
Matches 222; Conservative 0; Mismatches 146; Indels 6; Gaps 1;  
Qy 256 GGGCGGCCCTTCCCGAGGAGCTCCCTCCCTGCGAGCGGAGCTGGAGCGGCTGGG 315  
Db 455 GTGACGCGGCGAGCAGCGCGGGTTGGGGCGAGCTTGGCGGCCAGGTGGAGCGATTGGG 396  
Qy 316 GCCGAGCTGCGGGAGGAGCGCAAGGCCATCACAGATGTCTCGGGCTTCCAGCATGAG 375  
Db 395 GTGAGCTGCGAGCGAGCGCGGGTGGAGGAGCGCGGACAGCTTTGAGGGGAG 336  
Qy 376 CGGCTCGTGTGGAGGAGGAGAGAGTGTATTCAGTACCAGAAACAGCTGCAGCAG 435  
Db 335 CGGCTGGCTTGGCAGGAGAGAGGAGCAGGTGATCCGCTACAGAGCAGCTGCAGCAG 276  
Qy 436 AGCTAGTGGCCATGTATCCAGCGGAAACAGCGCTTGGAGAGGCCCTTGCAGCAGCTGGCA 495  
Db 275 AACTACATCCAGATGTATCCGCGCAACCGGCGAGCTAGAGCAGGAGCTGCAGCAGCTCAGC 216  
Qy 496 CGTGGGACAGCGCGGGAGCCCTTGGAGCTTGACCTGG-----AAGGGCTGACATC 549  
Db 215 CTGAGCTGGAGGCGCGGAGCTCGTGACCTGGGCTTGGCGAGCAGCGCCCTGTCATC 156  
Qy 550 CCCTACGAGGACATCATGACCTGAGATCTGAGGGGCTGCTGGGAGGCGAGTCTGGG 609  
Db 155 TGCCTGGAGGAGATCACTGCTACTGAGATCTAAGCCCTCAGCAACAGCTCTGTAGGA 96  
Qy 610 GACTTGGCACTGGG 623  
Db 95 GCTCTGCCAGAGGG 82

RESULT 15  
BQ694854  
LOCUS  
DEFINITION  
1000997 Human Fat Cell 5'-Stretch Plus cDNA Library Homo sapiens  
450 bp mRNA linear EST 15-JUL-2002  
CDNA 5', mRNA sequence.  
ACCESSION  
BQ694854  
VERSION  
BQ694854.1 GI:21820170  
KEYWORDS  
EST.

SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 450)  
AUTHORS  
Yang, R.-Z., Shuldiner, A. and Gong, D.-W.  
TITLE  
EST analysis of human adipose gene expression  
JOURNAL  
Unpublished (2002)  
COMMENT  
Contact: Gong Da-Wei  
Division of Endocrinology, Diabetes and Nutrition  
University of Maryland  
660 Redwood St, HH497, Baltimore, MD 21201, USA  
Tel: 410 706 1672  
Fax: 410 706 1622  
Email: dgong@medicine.umaryland.edu  
PCR Primers  
FORWARD: CTCGGGAAGCGCGCATTTGTGGT  
BACKWARD: AATACACTCATATAGGGGGAATTGG  
Seq primer: GTTGTACCCGGGAATTC.  
Location/Qualifiers  
1. .450  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Human Fat Cell 5'-Stretch Plus cDNA Library"  
/sex="Male and Female"  
/tissue\_type="Adipose"  
/note="Vector: lambdaTriplex"  
a 131 c 172 g 61 t

FEATURES  
source

Query Match 18.5%; Score 118.2; DB 14; Length 450;  
Best Local Similarity 61.2%; Pred. No. 1.1e-11;  
Matches 211; Conservative 0; Mismatches 128; Indels 6; Gaps 1;  
Qy 250 GACATGGGCGGCCACCTTCCCGAGGAGCTCCCTCCCTGCGAGCGGAGCTGGAGCGG 309  
Db 91 GCCAAAGTGCAGCGGCGAGCAGCGGGTTGGGGCGAGCTTGGCGGCCAGGTGGAGCGA 150  
Qy 310 CTGCGGCGCGAGCTGCGGGAGGAGCGCAAGGCCATCACAGATGTCTCTCGGGCTTCCAG 369  
Db 151 TTGCGGTGGAGCTGCAGCGGAGCGCGGGTGGAGGAGCGGAGCAGCTTTTTCAG 210  
Qy 370 CATGAGCGGCTCGTGTGGAGGAGGAGAGAGAGTGTATTCAGTACCAGAAACAGCTG 429  
Db 211 GGGAGCGGCTGGCTTGGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 270  
Qy 430 CAGCAGAGCTACGTGGCCATGTACAGCGGAGCAGCGCTTGGAGAGAGCGCTTGCAGCAG 489  
Db 271 CAGCACAACTACATCCAGATGTATCCGCGCAACCGGCGAGCTAGAGCAGGAGCTGCAGCAG 330  
Qy 490 CTGCGAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGACCTGG-----AAGGGGCT 543  
Db 331 CTAGCCTGGAGCTGGAGGCGCGGAGCTCGTGACCTGGGCTTGGCGAGCAGCGCCCC 390  
Qy 544 GACATCCCTTACGAGGACATCATAGCCACTGAGATCTGAGGGGCT 588  
Db 391 TGCACTCTGCTGGAGGAGATCACTGCTACTGAGATCTAGGGCCCT 435

Search completed: June 15, 2003, 03:41:57  
Job time : 894.672 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:39:06 ; Search time 4660.9 Seconds  
(without alignments)  
11183.052 Million cell updates/sec

Title: US-09-513-888C-3  
Perfect score: 1791  
Sequence: 1 atgggcagcgtcagtagcct.....tcatagcactgagatctga 1791

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1791	100.0	5492	9	AF123659	AF123659 Homo sapi
2	1643	91.7	1722	9	AF123658	AF123658 Homo sapi
3	1583	88.4	1692	9	AF123657	AF123657 Homo sapi
4	1427	79.7	1614	9	AF123655	AF123655 Homo sapi
5	1372.4	76.6	1515	9	AF123656	AF123656 Homo sapi
6	808.8	45.2	173264	2	AC025853	AC025853 Homo sapi
7	807.2	45.1	9108	9	AF123653	AF123653 Homo sapi
8	592.6	33.1	227884	2	AC114995	AC114995 Mus muscu
9	592.6	33.1	263546	2	AC099416	AC099416 Mus muscu
10	582.6	31.4	633	9	AF123654	AF123654 Homo sapi
11	475.4	26.5	191210	2	AC108987	AC108987 Rattus no
12	314	17.5	1935	9	AY029201	AY029201 Homo sapi
13	314	17.5	2275	9	HSM805394	AL834338 Homo sapi
14	314	17.5	2766	9	BC006212	BC006212 Homo sapi
15	314	17.5	5257	9	AB011124	AB011124 Homo sapi
16	314	17.5	5733	9	AB058716	AB058716 Homo sapi
17	310.8	17.4	2099	9	AB046013	AB046013 Macaca fa
18	284.8	15.9	2767	10	BC014695	BC014695 Mus muscu
19	278	15.5	1612	9	BC005855	BC005855 Homo sapi
20	274	15.3	191210	2	AC108987	AC108987 Rattus no
21	265.8	14.8	2416	10	RNO278801	AC278801 Rattus no
22	218.4	12.2	85530	2	AC125907	AC125907 Rattus no
23	197	11.0	181086	9	AL133215	AL133215 Human DNA
24	189	10.6	155662	2	AC112388	AC112388 Rattus no
25	186.4	10.4	272545	2	AC090533	AC090533 Mus muscu
26	184.6	10.3	180418	2	AC106674	AC106674 Rattus no
27	184.6	10.3	195673	2	AL807807	AL807807 Mus muscu
28	181	10.1	188210	2	AC116700	AC116700 Mus muscu
29	180.8	10.1	2308	9	AK097997	AK097997 Homo sapi
30	179.4	10.0	82400	10	AC090495	AC090495 Genomic s
31	168.8	9.4	103729	2	AC096620	Continuation (4 of
32	168.8	9.4	270700	2	AC105485	AC105485 Rattus no
33	168.6	9.4	167636	2	AC107097	AC107097 Rattus no
34	167.6	9.4	207683	2	AC098712	AC098712 Mus muscu
35	167.2	9.3	125856	9	HS1187M17	AF121891 Human DNA
36	166	9.3	231	9	AF123652	AF123652 Homo sapi
37	165.4	9.2	195690	10	AC098726	AC098726 Mus muscu
38	159.8	8.9	177035	2	AC099443	AC099443 Rattus no
39	157.4	8.8	141640	2	AC016204	AC016204 Homo sapi
40	157.4	8.8	180553	2	AC011767	AC011767 Homo sapi
41	156.8	8.8	198228	2	AC102097	AC102097 Mus muscu
42	156	8.7	227588	2	AC094571	AC094571 Rattus no
43	155.6	8.7	204259	2	AC110817	AC110817 Mus muscu
44	152.6	8.5	184865	2	AL807824	AL807824 Mus muscu
45	152.2	8.5	2508	9	ORAINVOL	M25312 Orang-utan

ALIGNMENTS

RESULT 1  
AF123659  
LOCUS AF123659 Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.  
DEFINITION  
ACCESSION AF123659  
VERSION AF123659.1 GI:4572475  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 5492)  
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,  
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.  
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,







```
Db 1021 GAGAGCGGAGCTCCGCGAGAGCTCGAGAGCTCATGAAGAGCAGAGCTCTGTGAG 1080
Qy 1081 ACCAAGCTCAGGTCTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGGAGAGACC 1140
Db 1081 ACCAAGCTCAGGTCTCTACGAGAGGAGAGACCAAGCTTCGGCCCGCGCTGGAGAGACC 1140
Qy 1141 CAGTGGAGGTGTGCGAGAGTCAAGAGTCAAGAGTCTCCCTCTCTGAAGCAGCAGCTGAAGAG 1200
Db 1141 CAGTGGAGGTGTGCGAGAGTCAAGAGTCAAGAGTCTCCCTCTCTGAAGCAGCAGCTGAAGAG 1200
Qy 1201 TCCAGAGCGAGGTGAAGCCAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAG 1260
Db 1201 TCCAGAGCGAGGTGAAGCCAGGCTAGCGAGATCTGGGTCTCAAGGCACAGCTGAAG 1260
Qy 1261 GACACGCGGCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTG 1320
Db 1261 GACACGCGGCGAAGCTGGAGGCGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCCCTG 1320
Qy 1321 CGCACCAGGCGCTGGAGCTGGAGTCTGTGAGATGAGCTGCAGCGCAGCAAGAACGAG 1380
Db 1321 CGCACCAGGCGCTGGAGCTGGAGTCTGTGAGATGAGCTGCAGCGCAGCAAGAACGAG 1380
Qy 1381 CGCGAGCTGCTGCGGAGAGGTGAACCTGTGAGCAGGAGCTGCAGGAGCTGCGGGCC 1440
Db 1381 CGCGAGCTGCTGCGGAGAGGTGAACCTGTGAGCAGGAGCTGCAGGAGCTGCGGGCC 1440
Qy 1441 CAGGCGGCGCTGGCGCGGACATGAGGCGCGCCACCTTCCCGAGGAGCTTCCCTGCCCTG 1500
Db 1441 CAGGCGGCGCTGGCGCGGACATGAGGCGCGCCACCTTCCCGAGGAGCTTCCCTGCCCTG 1500
Qy 1501 CAGCGGAGCTGGAGGCGCTGGGCGCGAGCTGGGAGGAGCGGCAAGGCCATGACGAG 1560
Db 1501 CAGCGGAGCTG----- 1512
Qy 1561 ATGCTCTCGGCTTCCAGCATGAGCGGCTCGTGTGAGAGGAGAGAGGAGAGGTGATT 1620
Db 1513 -----GAGCGGCTGTGTGAGAGGAGAGAGGAGAGGTGATT 1551
Qy 1621 CAGTACAGAAACAGCTCAGCAGAGCTACGTGGCCATGTACACGCGAAACAGCGCTG 1680
Db 1552 CAGTACAGAAACAGCTCAGCAGAGCTACGTGGCCATGTACACGCGAAACAGCGCTG 1611
Qy 1681 GAGAGGCGCTGCAGAGCTGCAGCTGGGACAGCGCGGAGCGCTTGGAGGTTGAC 1740
Db 1612 GAGAGGCGCTGCAGAGCTGCAGCTGGGACAGCGCGGAGCGCTTGGAGGTTGAC 1671
Qy 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1791
Db 1672 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGCCACTGAGATCTGA 1722
```

```
RESULT 3
AF123657
LOCUS AF123657
DEFINITION Homo sapiens clone G3611 FEZ1 (FEZ1) mRNA, linear PRI 07-APR-1999
complete cds.
ACCESSION AF123657
VERSION AF123657.1 GI:4572471
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1692)
Ishii.H., Baffa.R., Numata.S.I., Murakumo.Y., Rattan.S., Inoue.H.,
Mori.M., Fidanza.V., Alder.H. and Croce.C.M.
The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
99199287
PUBMED 10097140
REFERENCE 2 (bases 1 to 1692)
```

AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Ishii, H., Baffa, R., Numata, S. I., Murakumo, Y., Rattan, S., Inoue, H.,  
Mori, M., Fidanza, V., Alder, H. and Croce, C. M.  
Direct Submission  
Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer  
Institute, 2338 10th street, Philadelphia, PA 19107, USA  
Location/Qualifiers

1..1692  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22"  
/clone="G3611"  
1..1692  
/gene="FEZ1"  
1..1692  
/gene="FEZ1"  
/note="alternatively spliced"  
/codon\_start=1  
/product="FEZ1"  
/protein\_id="AAD23838.1"  
/db\_xref="GI:4572472"

gene  
CDS

BASE COUNT 380 a 522 c 554 g 236 t  
ORIGIN  
Query Match 88.4%; Score 1583; DB 9; Length 1692;  
Best Local Similarity 94.5%; Pred. No. 3.6e-214;  
Matches 1692; Conservative 0; Mismatches 0; Indels 99; Gaps 1;  
Qy 1 ATGGGCGAGCTCAGTACCTCATCTCGGCGCAGAGCTTCCACAGCAGCAGCTCGGGCT 60  
Db 1 ATGGGCGAGCTCAGTACCTCATCTCGGCGCAGAGCTTCCACAGCAGCAGCTCGGGCT 60  
Qy 61 TCGCAGTACAGCTGCGCAAGTCTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGG 120  
Db 61 TCGCAGTACAGCTGCGCAAGTCTCTCCACCTCAAGAGCTCAACCGGTATTCCGACGGG 120  
Qy 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAACGCAAGTCCAGCTCAAAATGGGC 180  
Db 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTCAACGCAAGTCCAGCTCAAAATGGGC 180  
Qy 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCCCGGGCTCCCATCACC 240  
Db 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGCAAGCCCGGGCTCCCATCACC 240  
Qy 241 GATTACAGGCACTGTCCAGCGGGATTAGGGGGCCAGGCTGGGGTGGACTTTGACCCG 300  
Db 241 GATTACAGGCACTGTCCAGCGGGATTAGGGGGCCAGGCTGGGGTGGACTTTGACCCG 300  
Qy 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGGCTCCGAGAAGGT 360  
Db 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGGCTCCGAGAAGGT 360  
Qy 361 GCAGTAGGGCCACAGCTTCAAGCTTGTGTCACGCTCAGGAGCAGCTCTGCATGCC 420  
Db 361 GCAGTAGGGCCACAGCTTCAAGCTTGTGTCACGCTCAGGAGCAGCTCTGCATGCC 420  
Qy 421 TCCCGGAGAGTGCAGCCAGCTGCACCCGCCCTTCCAGACAGCCCAAGGAGCAG 480  
Db 421 TCCCGGAGAGTGCAGCCAGCTGCACCCGCCCTTCCAGACAGCCCAAGGAGCAG 480  
Qy 481 GAGCTGAAGCTGCGCTGTGCTCTGGGCGCTGTGAGACTCCGCGCGGAACTCCATGTC 540  
Db 481 GAGCTGAAGCTGCGCTGTGCTCTGGGCGCTGTGAGACTCCGCGCGGAACTCCATGTC 540

Qy	1621	CAGTACCAAGAACAGCTGCAGACAGACTACGTGGCCATGTACCAAGCGGAACCAAGCCGCTG	1588
Db	1522	CAGTACCAAGAACAGCTGCAGACAGACTACGTGGCCATGTACCAAGCGGAACCAAGCCGCTG	1581
Qy	1681	GAGAAGGCCCTTCAGACAGCTGGCACGCTGGGGACAGCGCCGGGGAGCCCTTGGAGGTTGAC	1740
Db	1582	GAGAAGGCCCTTCAGACAGCTGGCACGCTGGGGACAGCGCCGGGGAGCCCTTGGAGGTTGAC	1641
Qy	1741	CTGGAAGGGGCTGACATCCCTACGAGGACATCATAGCCACCTGAGATCTGA	1791
Db	1642	CTGGAAGGGGCTGACATCCCTACGAGGACATCATAGCCACCTGAGATCTGA	1692
RESULT 4			
AF123655 1614 bp mRNA linear PRI 07-APR-1999			
LOCUS Homo sapiens clone T8D145M4 FEZ1 (FEZ1) mRNA, alternatively			
DEFINITION spliced, complete cds.			
ACCESSION AF123655			
VERSION AF123655.1			
KEYWORDS GI:4572467			
SOURCE Homo sapiens.			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE 1 (bases 1 to 1614)			
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,			
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.			
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,			
and its expression is altered in multiple human tumors			
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)			
JOURNAL 99199287			
MEDLINE 10097140			
PUBMED			
REFERENCE 2 (bases 1 to 1614)			
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,			
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.			
TITLE Direct Submision			
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer			
Institute, 2338 10th street, Philadelphia, PA 19107, USA			
FEATURES			
source			
1..1614			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/chromosomes="8"			
/map="8p22"			
/clone="T8D145M4"			
1..1614			
/gene="FEZ1"			
1..1614			
/gene="FEZ1"			
/note="alternatively spliced"			
/codon_start=1			
/product="FEZ1"			
/protein_id="AAD24836.1"			
/db_xref="GI:4572468"			
/translation="MGVSLSLISGHSHFKHCRASQYKLRKSHLKLNRYSDGLLRF			
GFSQDNGHCKSSKMGKSDFFYIKVQSKARGSHHPDYIALSSGDLGGQAGVDFDST			
PPLKLPFNSGLQEMSEKCAVRPTAKPVLPRSGAILHSSPESASHQLHPAPDPKPEQ			
ELKPLGCSALSDSGRNSMSLPHTSTSSSYQLDRLVTPVGTSPFGGSAHNITQGI			
LDQNMMSLUKALSDFDGSGLKSHNKAADGPCSVSPISITDECSIQELQKILLERGA			
LQKQRSYSEKELASLAYEERPRCRDELEGPFGKGNLKAQSKQSQAQOVLHLQ			
VLQQLQKQRLQOELLSLMKEQDLLETKLRSYERKTSFPALEETQWMEVCCKSGSIS			
VLQQLQKESQTVNKAESIILGLKAQLKDTRCKLEGLRTODLEGALTKGLELEVC			
ENELORKNEALLREKHERLVWKEEKEKVIQYQKLQSQSYVAMVQRNORLEKALQOL			
AKGDSAGEPLEVLDLSGADIPYEDIIATEI"			
BASE COUNT 367 a 500 c 522 g 225 t			
ORIGIN			
Query Match 79.7%; Score 1427; DB 9; Length 1614;			
Best Local Similarity 90.1%; Pred. No. 3.7e-192;			
Matches 1614; Conservative 0; Mismatches 0; Indels 177; Gaps 1;			
Qy	1	ATGGGCAGCGCTCAGTAGCCTCATCTCCGCCACAGCTTCCACAGCAACGACCTGCCCGGCT	60



FEATURES	Location/Qualifiers	
	1..1515	
source	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
gene	/chromosome="8"	
	/map="8p22"	
CDS	/clone="D14"	
	1..1515	
note="alternatively spliced"	/gene="FEZ1"	
	/note="alternatively spliced"	
protein id="AAD23837.1"	/codon_start=1	
	/product="FEZ1"	
trnlat="WGSVSLISGHSFHSKCRASOYKLRKSSHLKLNRYSDGLLR	/db_xref="GI:4572470"	
	/trnlat="WGSVSLISGHSFHSKCRASOYKLRKSSHLKLNRYSDGLLR	
GFSDDSGHGKSSXMGKSEDFYIKVQKARGSHHPYATALLSDGLGGQAGVDFDPS	PKLMPSPNQJEMSEKGAVRPTAFKVPVLRSGAILHSSPESASHQHPAPDPKPEQ	
	ELKPLGSLGALSGRMSLSLTHSTSSYOLDPLVTPVGFTRFPGSAHNTQGV	
LQDNMMSLKALSFSDGSLGHNKADKPGSCVRSPISTDECSIQLEQKLLEREGA	LQKQRFEEELASSLAYERPRRCRDELEPGPKGNKLLQASQSORAQVHLQ	
	VLQIQERKROLRLESLMKPDLLLETKLRSYEREKTSFGPALBETOWEVCOKSGEIS	
LLKQOLKESQTEVNAKASEILGLKQKDTFGKLEGLRPTDLEGALRTKGLRLEV	ENELQOSVAMYQNRLEKALQQLARGDSAGEFLEVDLEGADIPYEDIATET	
	336 a 483 c 482 g 214 t	
BASE COUNT	336 a 483 c 482 g 214 t	
ORIGIN		
Query Match	76.6%; Score 1372.4; DB 9; Length 1515;	
	Best Local Similarity 95.4%; Pred. No. 1.9e-184;	
Matches 1425; Conservative 0; Mismatches 66; Indels 3; Gaps 1;		
Qy	1	ATGGGACGCTCAGTAGCTCATCTCCGGCCACACAGCTTCCACAGCAAGCACTGCCGGGCT 60
Db	1	ATGGGACGCTCAGTAGCTCATCTCCGGCCACACAGCTTCCACAGCAAGCACTGCCGGGCT 60
Qy	61	TCCGAGTACAAGTGGCGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120
Db	61	TCCGAGTACAAGTGGCGAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120
Qy	121	CTGCTGAGGTTGGCTTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGGC 180
Db	121	CTGCTGAGGTTGGCTTCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAATGGGC 180
Qy	181	AAGAGCGAAGACTTCTTACATCAAGGTACAGCAAGGCTCCAGCTCCAAATGGGC 240
Db	181	AAGAGCGAAGACTTCTTACATCAAGGTACAGCAAGGCTCCAGCTCCAAATGGGC 240
Qy	241	GATTACAGGCACTGTCAGCGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 300
Db	241	GATTACAGGCACTGTCAGCGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 300
Qy	301	TCCACACCCCGGCTATGCTTCCATCAGCTAGCAAGGCTCCAGCTCCGAGAGGCT 360
Db	301	TCCACACCCCGGCTATGCTTCCATCAGCTAGCAAGGCTCCAGCTCCGAGAGGCT 360
Qy	361	GCAGTGAGGCCACAGACCTTCAAGCTGTGTGTCACCGGTACAGGCTCCGACCTCC 420
Db	361	GCAGTGAGGCCACAGACCTTCAAGCTGTGTGTCACCGGTACAGGCTCCGACCTCC 420
Qy	421	TCCCGGAGAGTCCAGCCACAGCTGTCACCCCGGCTCCAGCAAGCCCAAGGAGCAG 480
Db	421	TCCCGGAGAGTCCAGCCACAGCTGTCACCCCGGCTCCAGCAAGCCCAAGGAGCAG 480
Qy	481	GAGCTGAAGCTGGCTGCTGCTGGGGCTGTGTCAGACTCCGGCCGGAATCCATGTCC 540
Db	481	GAGCTGAAGCTGGCTGCTGCTGGGGCTGTGTCAGACTCCGGCCGGAATCCATGTCC 540
Qy	541	AGCTGCCACACACAGCAGCAGCTACAGCTGGAGCCCGCTGGTCACACCCGCTG 600
Db	541	AGCTGCCACACACAGCAGCAGCTACAGCTGGAGCCCGCTGGTCACACCCGCTG 600
Qy	601	GGACCCACAAGCGGTTTTGGGGGCTCCGGCCCAACAACATCACCCAGGGGATCGTCTCCAG 660

Db	601	GGACCCACAAGCGGTTTTGGGGGCTCCGGCCCAACAACATCACCCAGGGGATCGTCTCCAG 660
Qy	661	GACAGCAACATGATGAGCTTGAAGGCTGTGCTTCTCCGAGCGAGGTAGCAAGCTGGGC 720
Db	661	GACAGCAACATGATGAGCTTGAAGGCTGTGCTTCTCCGAGCGAGGTAGCAAGCTGGGC 720
Qy	721	CATTCGAACAAAGCAGACAAGGGCCCTCGTGTGTCTCCGCTCCCGCTCCAGGACGAG 780
Db	721	CATTCGAACAAAGCAGACAAGGGCCCTCGTGTGTCTCCGCTCCCGCTCCAGGACGAG 780
Qy	781	TGCAGCATCCAGGAGCTGGAGCAGAAGCTGTGTGAGAGGGAGGGCGCTCCAGAGCTG 840
Db	781	TGCAGCATCCAGGAGCTGGAGCAGAAGCTGTGTGAGAGGGAGGGCGCTCCAGAGCTG 840
Qy	841	CAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCTTGGCTTACGAGAGCGCGCGGG 900
Db	841	CAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCTTGGCTTACGAGAGCGCGCGGG 900
Qy	901	CGCTGCAGGAGCAGCTGGAGGGCCCGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCC 960
Db	901	CGCTGCAGGAGCAGCTGGAGGGCCCGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCC 960
Qy	961	TCCGAGAGAGCCAGCGCGGCGCAGAGTCTGCTGCACTGCAAGTCTGCAAGTCTCAGCAG 1020
Db	961	TCCGAGAGAGCCAGCGCGGCGCAGAGTCTGCTGCACTGCAAGTCTGCAAGTCTCAGCAG 1020
Qy	1021	GAGAAGCGGAGCTCCCGGAGGAGCTCGAGAGCTCATGAAGGAGCAGGAGCTGCTGGAG 1080
Db	1021	GAGAAGCGGAGCTCCCGGAGGAGCTCGAGAGCTCATGAAGGAGCAGGAGCTGCTGGAG 1080
Qy	1081	ACCAAGCTCAGGTTCTACGAGAGGAGAAGACCAAGCTTCCGCCCGCGCTGGAGGAGACC 1140
Db	1081	ACCAAGCTCAGGTTCTACGAGAGGAGAAGACCAAGCTTCCGCCCGCGCTGGAGGAGACC 1140
Qy	1141	CAGTGGAGGTGTCCAGAGTCAAGGCGAGATCTCCCTCTCTGAAGCAGCAGCTGAAGAG 1200
Db	1141	CAGTGGAGGTGTCCAGAGTCAAGGCGAGATCTCCCTCTCTGAAGCAGCAGCTGAAGAG 1200
Qy	1201	TCCGAGAGCGAGGTGAACGCAAGGCTAGCGAGATCTGGGTCTCAAGGACACAGCTGAAG 1260
Db	1201	TCCGAGAGCGAGGTGAACGCAAGGCTAGCGAGATCTGGGTCTCAAGGACACAGCTGAAG 1260
Qy	1261	GACACGCGGGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTG 1320
Db	1261	GACACGCGGGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTG 1320
Qy	1321	CGACACAGGGCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGC---GCAAGAGAAC 1377
Db	1321	CGACACAGGGCTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGC---GCAAGAGAAC 1377
Qy	1378	GAGCGGAGCTGTGCGGAGAGGTGAACCTGCTGGAGGAGGAGCTGACAGGAGCTGCGG 1437
Db	1381	ATGTACACAGCGAAGACCGGCTGGAGAGGCCCTGACAGAGCTGGACGTGGGAGACG 1440
Qy	1438	GCCAGGCGGCTGGCGGACATGGGGCGGCCACCTTCCCGGAGGAGCTC 1491
Db	1441	GCCGGGAGGCTTTGGAGTTGACCTGGAAGGGCTGACATCCCTACGAGGAC 1494

RESULT 6  
AC025853  
LOCUS  
DEFINITION Homo sapiens chromosome 8 clone RP11-353K12 map 8, \*\*\* SEQUENCING  
IN PROGRESS \*\*\*, 3 unordered pieces.  
AC025853  
VERSION AC025853.13 GI:21431202  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 173264)

AUTHORS  
JOURNAL  
REFERENCE  
AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 8, clone RP11-353K12  
Unpublished  
2 (bases 1 to 173264)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bada,F.,  
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marguis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,  
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenka,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 173264)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,  
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,  
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,  
Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N.,  
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J.,  
Meneus,L., Mihova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C.,  
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,  
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,  
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,  
Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,  
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (17-JUN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 17, 2002 this sequence version replaced gi:21321864.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L7454  
Center clone name: 353\_K\_12  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 39461: contig of 39461 bp in length  
\* 39462 39561: gap of 100 bp  
\* 39562 108347: contig of 68786 bp in length  
\* 108348 108447: gap of 100 bp  
\* 108448 173264: contig of 64817 bp in length.  
FEATURES  
Location/Qualifiers  
source  
1..173264  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-353K12"  
/clone\_lib="RP11-11 Human Male BAC"  
BASE COUNT 47355 a 38864 c 39235 g 47493 t 317 others  
ORIGIN  
Query Match 45.2%; Score 808.8; DB 2; Length 173264;  
Best Local Similarity 99.8%; Pred. No. 2.7e-105;  
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 345 GGGCTCCGAGAGGGTGCAGTGGAGGCCACACAGCCTTCAAGCCTGTGTGTCACACGGTCAGG 404  
Db 23928 GGGCTCCGAGAGGGTGCAGTGGAGGCCACACAGCCTTCAAGCCTGTGTGTCACACGGTCAGG 23987  
Qy 405 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCACCACAGCTGCACCCGCCCTCCAGA 464  
Db 23988 AGCCATCTGCATCTCTCCCGGAGAGTGCAGCACCACAGCTGCACCCGCCCTCCAGA 24047  
Qy 465 CAAGCCCAAGAGGAGGAGTGAAGCCTGGCTGTGTCTGGGGCGGTGTGTCAGACTCCGG 524  
Db 24048 CAAGCCCAAGAGGAGGAGTGAAGCCTGGCTGTGTCTGGGGCGGTGTGTCAGACTCCGG 24107  
Qy 525 CCGGAATCTCATGTGTCAGCTGGCCACACAGCACCAGCAGCAGTACACAGCTGGACCC 584  
Db 24108 CCGGAATCTCATGTGTCAGCTGGCCACACAGCACCAGCAGCAGTACACAGCTGGACCC 24167  
Qy 585 GCTGGTCACACCCGTGGAGCCCAAGCCGTTTTTGGGGGCTCCGCCACACATCACCACA 644  
Db 24168 GCTGGTCACACCCGTGGAGCCCAAGCCGTTTTTGGGGGCTCCGCCACACATCACCACA 24227  
Qy 645 GGGCATCTGCTCTCCAGACACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 704  
Db 24228 GGGCATCTGCTCTCCAGACACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 24287  
Qy 705 AGGTAGCAAGTGGGCGCCTCGAACAGAGCAGACAAAGGCGCCCTCGTGTGTGTCGCTCC 764  
Db 24288 AGGTAGCAAGTGGGCGCCTCGAACAGAGCAGACAAAGGCGCCCTCGTGTGTGTCGCTCC 24347  
Qy 765 CATCTCCAGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTTGGAGGGGAGGG 824  
Db 24348 CATCTCCAGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTTGGAGGGGAGGG 24407  
Qy 825 CGCCCTTCAGAAAGTGCAGCGCAGCTTTTGAAGGAGAGAGCTTCGCTCCAGCCTGGCCTA 884  
Db 24408 CGCCCTTCAGAAAGTGCAGCGCAGCTTTTGAAGGAGAGAGCTTCGCTCCAGCCTGGCCTA 24467  
Qy 885 CGAGGAGCGGCGCGCGCTGTCAGGAGCAGAGCTGGAGGGCCCGAGGCCCAAGGCGGCAA 944  
Db 24468 CGAGGAGCGGCGCGCGCTGTCAGGAGCAGAGCTGGAGGGCCCGAGGCCCAAGGCGGCAA 24527  
Qy 945 CAAGCTCAAGCAGGCTCCGAGAGAGCCAGCGCGCCGAGCAGTCTGTCACCTGCAGGT 1004  
Db 24528 CAAGCTCAAGCAGGCTCCGAGAGAGCCAGCGCGCCGAGCAGTCTGTCACCTGCAGGT 24587  
Qy 1005 ACTCAGCTTCAGCAGGAGAGCGGCGAGCTCCGCGAGGAGCTTCGAGAGCTTCATGAAGGA 1064  
Db 24588 ACTCAGCTTCAGCAGGAGAGCGGCGAGCTCCGCGAGGAGCTTCGAGAGCTTCATGAAGGA 24647  
Qy 1065 GCAGGACCTGCTGAGACCAAGCTCAGGTCTCTACGAGGGAGAGACCAAGCTTCGGGCC 1124  
Db 24648 GCAGGACCTGCTGAGACCAAGCTCAGGTCTCTACGAGGGAGAGACCAAGCTTCGGGCC 24707

AC114995	227884 bp	DNA	linear	HTG 26-JUN-2002
LOCUS				
DEFINITION	Mus musculus clone RP23-171P5, WORKING DRAFT SEQUENCE, 5 ordered			



Ginde S., Gord S., Goyette M., Graham L., Grand-Pierre N., Hagois B., Horton L., Hulme W., Iliev I., Johnson R., Jones C., Kanat A., Karatas A., Kells C., LaRocque K., Lamazares R., Landers T., Lehoczyk J., Levine R., Lindblad-Toh K., Liu G., MacCarthy C., MacDonald P., Major J., Marquis N., Matthews C., McEwan P., McKernan K., Meldrim J., Meneses L., Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R., Norbu C., Norman C.H., O'Connor T., O'Donnell P., O'Neil D., Oliver J., Peterson K., Phunkhang P., Pierre N., Pollara V., Raymond C., Retta R., Rieback M., Riley R., Rise C., Rogov P., Roman J., Rosetti M., Roy A., Santos R., Schauer S., Schuback R., Seaman S., Severy P., Spencer B., Stange-Thomann N., Stevanus N., Subramanian A., Talamas J., Tesfaye S., Theodore J., Topham K., Travers M., Travis N., Trigilio J., Vassiliev H., Viel R., Vo A., Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zembek L., Zimmer A. and Zody M.

Direct Submission  
Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 227884)

Birken B., Linton L., Nusbaum C., Lander E., Ali A., Allen N., Anderson S., Barna N., Bastien V., Bloom T., Boguski L., Boukhgalter B., Brown A., Camarata J., Campotiano A., Chang J., Chazaro B., Choepel Y., Collangelo M., Collins S., Collamore A., Cook A., Cooke P., DeArillano K., Dewar K., Diaz J.S., Dodge S., Fato S., Ferreira P., Fitzgerald M., FitzHugh W., Gage D., Galagan J., Gardyna S., Ginde S., Gord S., Goyette M., Graham L., Grand-Pierre N., Hagois B., Horton L., Hulme W., Iliev I., Johnson R., Jones C., Kamat A., Karatas A., Kells C., Lakocque K., Lamazares R., Landers T., Lehoczyk J., Levine R., Lindblad-Toh K., Liu G., MacLean C., MacDonald P., Major J., Marquis N., Matthews C., McCarthy M., McEwan P., McKernan K., Meldrim J., Meneses L., Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R., Norbu C., Norman C.H., O'Connor T., O'Donnell P., O'Neil D., Oliver J., Peterson K., Phunkhang P., Pierre N., Pollara V., Raymond C., Retta R., Rieback M., Riley R., Rise C., Rogov P., Roman J., Rosetti M., Roy A., Santos R., Schauer S., Schuback R., Seaman S., Severy P., Spencer B., Stange-Thomann N., Stevanus N., Subramanian A., Talamas J., Tesfaye S., Theodore J., Topham K., Travers M., Travis N., Trigilio J., Vassiliev H., Viel R., Vo A., Wilson B., Wu X., Wyman D., Ye W.J., Young G., Zainoun J., Zembek L., Zimmer A. and Zody M.

Direct Submission  
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 26, 2002 this sequence version replaced gi:21535965.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIER

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submission@genome.wi.mit.edu](mailto:sequence_submission@genome.wi.mit.edu)

----- Project Information

Center project name: L21473

Center clone name: 171\_P5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 225162 bases at least Q40

Consensus quality: 226535 bases at least Q30

Consensus quality: 227117 bases at least Q20

Insert size: 225000; agarose-EP

Insert size: 227484; sum-of-contigs

Quality coverage: 8.6 in Q20 bases; agarose-EP

Quality coverage: 8.6 in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently

\* consists of 5 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces

\* is believed to be correct as given, however the sizes

\* of the gaps between them are based on estimates that have  
\* provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 301: contig of 301 bp in length

\* 302 401: gap of 100 bp

\* 402 5831: contig of 5430 bp in length

\* 5832 5931: gap of 100 bp

\* 5932 173029: contig of 167098 bp in length

\* 173030 173129: gap of 100 bp

\* 173130 210061: contig of 36932 bp in length

\* 210062 210161: gap of 100 bp

\* 210162 227884: contig of 17723 bp in length.

#### FEATURES

source

1..227884

/organism="Mus musculus"

/db xref="taxon:10090"

/clone="RP23-171P5"

/clone\_lib="RPCI-23 Female Mouse BAC"

misc\_feature

1..301

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left

402..5831

/note="assembly\_fragment"

5932..173029

/note="assembly\_fragment"

173130..210061

/note="assembly\_fragment"

210162..227884

/note="assembly\_fragment"

clone\_end:T7

vector\_side:right

BASE COUNT 65784 a 48825 c 49902 g 62967 t 402 others

#### ORIGIN

Query Match 33.1%; Score 592.6; DB 2; Length 227884;

Best Local Similarity 84.0%; Pred.No. 8e-75; Indels 6; Gaps 1;

Matches 683; Conservative 0; Mismatch 124;

Qy 345 GGGTCGCGAAGGGTGCAGTGGCGCCACAGCCTTCAAGCCTGTGTGCGCACGGTCAGG 404

Db 84371 GAGTTACAGCAAGGGGCGAGTGAAGGCCACTGCTTCAAGCCTGTGTGCGCACGGTCAGG 84430

Qy 405 AGCATCTGTGACTCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCGCCCTCCAGA 464

Db 84431 AGCATCTGTGACTCTCTCCCGGAGAGTGCAGCCACAGCTGCACCCGCCCTCCAGA 84490

Qy 465 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTGTCTTGGGGCGCTGTGAGACTCCGG 524

Db 84491 TAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTGTCTTGGGGCGCTGTGAGACTCCGG 84550

Qy 525 CCGGAATCCATGTCCAGCCTGCCACACACAGCAGCAGCAGCAGCTACAGCTGGAGCC 584

Db 84551 CCGGAATCCATGTCCAGCCTGCCACACACAGCAGCAGCAGCAGCTACAGCTGGAGCC 84610

Qy 585 GCTGGTACACCCCGTGGGACCCCAAGCCGTTTTGGGGGCTCCGCCACACATCACCA 644

Db 84611 TCTGGTACACCCCGTGGGACCCCAAGCCGTTTTGGGGGCTCCGCCACACATCACCA 84670

Qy 645 GGGCATCTCTCCAGCAGCAGCAATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 704

Db 84671 AGGCATCTCTCCAGCAGCAGCAATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 84730

Qy 705 AGGTAGCAAGCTGGGCGCTCGAACAAGGCGCCCTCGTGTGTCCGCTCCCC 764

Db 84731 CGGCAGCAAGCTGGGCGCTCGAACAAGGCGCCCTCGTGTGTCCGCTCCCC 84790

Qy 765 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTCGAGCAGAGCTGTGGAGAGGGAGG 824

Db 84791 ACTCTCCACGGAGTGCAGCATCCAGGAGCTCGAGCAGAGCTGTGGAGAGGGAGG 84850



Sequencing vector: M13, 32%  
 Sequencing vector: plasmid, 68%  
 Chemistry: Dye-primer ET, 0% of reads  
 Chemistry: Dye-terminator, Big Dye, 100% of reads  
 Assembly program: Phrap, version 0.990319  
 Consensus quality: 278310 bases at least Q40  
 Consensus quality: 282938 bases at least Q30  
 Consensus quality: 286149 bases at least Q20  
 Insert size: 250000; agarose-1p  
 Insert size: 387570; sum-of-contigs  
 Quality coverage: 25.74 in Q20 bases; agarose-1p

Query Match	33.1%;	Score 592.6;	DB 2;	Length 263546;
Best Local Similarity	84.0%;	Prod. No. 7.8e-75;		
Matches 683;	Conservative 0;	Mismatches 124;	Indels 6;	G
QY	345	GGGCTCCGAGAAGGGTGCATGAGGGCCCA	CAGCGCTTCAAGCGCTGTGCTGCCACGGT	
Db	250515	GAGTTACAGACAAGGGGCAGTAA	GGCCCACTGCCTTCAGAGCTGTGCTGCCACGGT	
QY	405	AGCCATCTGCATCTCTCCCGGAGAGTGC	CAGCGCACACAGCTTGCAACCCCGGCCCTCT	
Db	250455	AGCCATCTCCACCTTTCACTCTGAGAGT	ACCAAGCACCACCTTCATCCCATGCGCTCT	
QY	465	CAAGCCCCAAGGAGCAGGAGCTGAAGCG	CTGGCGCTGTGCTCTGGGGCGCTGTGCAGCT	
Db	250395	TAAGCCCAAGGAGCAGGAGCTGAAGCC	AGGCGCTGTGCTCTCGGGGCACCTGTCTGACT	
QY	525	CCGGNATCTCATGTCTAGCGCTGCCCA	CACACAGCACAGCAGCAGCTACAGAGCTGG	
Db	250335	CCGGAATCTCATGTCTAGCTGTGCCAC	GCATAGCACCAACAGCAGCTTACCAAGCTGG	
QY	585	GCTGTCACACCCCGTGGGACCCACAAG	CGCGCTTTTGGGGGCTCCCGGCCACCAACATCA	
Db	250275	TCTGTCACCCCGTGGGGCTTACAGCG	TTTTGGGGGCTTTCAGTCTCACACATCTCA	

```
QY 645 GGGCATCTCTCCAGGACAGCAATGATGAGCCTGAAGCCTCTGCTCTTCGACGG 704
Db 250215 AGGCATCATCTTTCAGGACAGTAATATGATGAGCCTGAAGCCTCTGCTCTTCGATGG 250156

QY 705 AGGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAGGCGCCCTCGTGTGCGCTCCCC 764
Db 250155 CGGACGCAAGCTGGCTCACCCAGGCAAGCAGATGAAGGCGCCTCTGTGTGCGTTCGCC 250096

QY 765 CATCTCCACGACGAGTGCAGATCCAGGAGCTGGAGCAGAGCTGTTGGAGAGGAGGG 824
Db 250095 ACTCTCCACGAGTGTGACCATCCAGGAGCTGGAGCAGAGCTGTGTCAGAGGAGAC 250036

QY 825 CGCCCTCAGAGCTGCAGCGAGCTTGGAGAGAGGAGCTGCTCCAGCCTGCGCTA 884
Db 250035 TGCACTCAGAGCTTACAGCGCAGTTTCGATGAGAAGGAGTTGCTCTGCGCAACCTT 249976

QY 885 CGAGAGCGGCGCGCGCTGCGAGGAGCAGCTGGAGGCGCGGAGCCCAAGGCGGCA 944
Db 249975 TGAGAGCGGCGCACCGCGCACTAGAGAGCACTGGAGTGCCTTGGAGCCTTA-----GAG 249922

QY 945 CAAGCTCAAGCAGGCTCCGAGAAGCCAGCGCGCGCAGCAGGTCTGCACTGCGAGGT 1004
Db 249921 CAAGCTGAAGCACCCTCACAGAAGCCAGCGCACAGCAGGTGCTCCAGCTGCGAGT 249862

QY 1005 ACTGAGCTTCAGCAGGAGAGCGCGAGCTCGGAGGAGCTCGAGAGCCTCATGAAGGA 1064
Db 249861 GTTGAGTTGACAGGAGGAGAACCGCGCAGCTAAGCGCAGGCTAGAGAGCCTTAATGAAGGA 249802

QY 1065 GCAGGACCTGTGGAGACCAAGCTCAGGTCTTACAGAGGAGAGAGCAGCTTCGGGCC 1124
Db 249801 GCAGGACCTGTGGAGACCAAGCTCAGGTCTTACAGAGGAGAGAGCAGCTTCGGGCC 249742

QY 1125 CGCGCTGGAGGAGAGCCAGTGGGAGGTGTGCA 1157
Db 249741 TGCACTGGAGGAGAGCCCAATGGGAGGTAAAGCA 249709

RESULT 10
AF123654
LOCUS
DEFINITION Homo sapiens clone E264162 FEZ1 (FEZ1) mRNA linear PRI 07-APR-1999
complete cds.
ACCESSION AF123654
VERSION AF123654.1 GI:4572465
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
and its expression is altered in multiple human tumors
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
MEDLINE 9919287
PUBMED 10097140
REFERENCE 2 (bases 1 to 633)
AUTHORS Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
Institute, 2335 10th street, Philadelphia, PA 19107, USA
FEATURES
Location/Qualifiers
1..633
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22"
/clone="E264162"
1..633
/gene="FEZ1"
```

```
CDS
1..633
/gene="FEZ1"
/notes="alternatively spliced"
/codon_start=1
/product="FEZ1"
/protein_id="AAD23835.1"
/db_xref="GI:4572466"
/translation="MGSVSLISGHSFHSKHCASOYKLRKSSHLKLNRYSDGLLRF
GFQDSGHGKSSSMKMSKSEDFYIKYSQAKRGSHHDPYATLSGDLGGQAGVDFDPT
PPKLMPFNSQLEMGSEKGAIRAFKPLFRSGAILHSSPESASHQHPAPDPKPEQ
ELKPGLCSGSLSDGRNSSLPSFTHPSRPLEVDLEGADIPYEDIIATEI"
BASE COUNT 136 a 217 c 175 g 105 t
ORIGIN
Query Match 31.4%; Score 562.6; DB 9; Length 633;
Best Local Similarity 97.6%; Pred. No. 3.8e-70;
Matches 571; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 1 ATGGGCGAGCGTCAGTAGCTTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60
Db 1 ATGGGCGAGCGTCAGTAGCTTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60
QY 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120
Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120
QY 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGGCAGTCCAGTCCAAAATGGC 180
Db 121 CTGCTGAGTTTGGCTTCTCCAGGACTCCGGTCACGGCAGTCCAGTCCAAAATGGC 180
QY 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGAAAGCCCGGGCTCCCATCACC 240
Db 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGAAAGCCCGGGCTCCCATCACC 240
QY 241 GATTACACGCACTGTCCAGCGGGATTTAGGGGGCCAGGCTGGGGTGGAATTGACCCG 300
Db 241 GATTACACGCACTGTCCAGCGGGATTTAGGGGGCCAGGCTGGGGTGGAATTGACCCG 300
QY 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCAACTCAGTCTAGTAAGTGGGCTCCGAGAGG 360
Db 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCAACTCAGTCTAGTAAGTGGGCTCCGAGAGG 360
QY 361 GCAGTGAAGGCGCCACAGCCTTCAAGCCTGTGTGTCACGCTCAGGAGCCTCTGCACTCC 420
Db 361 GCAGTGAAGGCGCCACAGCCTTCAAGCCTGTGTGTCACGCTCAGGAGCCTCTGCACTCC 420
QY 421 TCCCGGAGAGTGCAGCCACAGCTGACACCCCGCCCTCCAGACACCCCAAGAGGAG 480
Db 421 TCCCGGAGAGTGCAGCCACAGCTGACACCCCGCCCTCCAGACACCCCAAGAGGAG 480
QY 481 GAGCTGAAGCTGCGCTGTGCTCTGGGGCGCTGTGTCAGACTCCGGCGGAACCTCCATGTC 540
Db 481 GAGCTGAAGCTGCGCTGTGCTCTGGGGCGCTGTGTCAGACTCCGGCGGAACCTCCATGTC 540
QY 541 AGCTGCGCCACACAGCAGCAGCAGCTTACAGCTGAGGAGCCCG 585
Db 541 AGCTGCGCCACACAGCAGCAGCAGCAGCTTGGAGGTTGACCTG 585

RESULT 11
AC108987/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-115K1, *** SEQUENCING IN PROGRESS
***, 78 unordered pieces.
ACCESSION AC108987
VERSION AC108987.3 GI:21737647
KEYWORDS HTG; HTGS PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 191210)
```

AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Albraccio, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barabara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Huly, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lorado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Kapue, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 191210)  
Worley, K.C.

Submitted (03-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 191210)  
Worley, K.C.

Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18846600.

COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information -----  
Center project name: GPMW  
Center clone name: CH230-115K1  
----- Summary Statistics -----  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 115719 bases at least Q40  
Consensus quality: 122885 bases at least Q30  
Consensus quality: 129025 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently consists of 78 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1041: contig of 1041 bp in length  
1042: gap of unknown length  
1142: contig of 1121 bp in length  
2262: contig of 1121 bp in length  
2362: gap of unknown length  
3538: contig of 1176 bp in length  
3638: gap of unknown length  
4775: contig of 1137 bp in length  
4875: gap of unknown length  
6050: contig of 1175 bp in length  
6051: gap of unknown length  
7660: contig of 1510 bp in length  
7661: gap of unknown length  
7761: contig of 1571 bp in length  
9332: gap of unknown length  
9432: contig of 1596 bp in length  
11027: gap of unknown length  
11127: contig of 1642 bp in length  
12769: gap of unknown length  
12869: gap of unknown length  
14092: contig of 1223 bp in length  
14192: gap of unknown length  
15228: contig of 1036 bp in length  
15328: gap of unknown length  
16709: contig of 1381 bp in length  
16809: gap of unknown length  
18195: contig of 1386 bp in length  
18295: gap of unknown length  
18412: contig of 1117 bp in length  
19512: gap of unknown length  
20959: contig of 1447 bp in length  
20960: gap of unknown length  
22161: contig of 1102 bp in length  
22261: gap of unknown length  
23587: contig of 1326 bp in length  
23687: gap of unknown length  
25231: contig of 1544 bp in length  
25331: gap of unknown length  
27135: contig of 1804 bp in length  
27235: gap of unknown length  
28175: contig of 1940 bp in length  
29275: gap of unknown length  
30846: contig of 1571 bp in length  
30847: gap of unknown length  
32090: contig of 1144 bp in length  
32190: gap of unknown length  
34039: contig of 1849 bp in length  
34139: gap of unknown length  
35408: contig of 1269 bp in length  
35508: gap of unknown length  
37018: contig of 1510 bp in length  
37118: gap of unknown length  
38260: contig of 1142 bp in length  
38360: gap of unknown length  
40083: contig of 1723 bp in length  
40183: gap of unknown length  
41200: contig of 1017 bp in length  
41300: gap of unknown length  
43365: contig of 2065 bp in length  
43465: gap of unknown length  
45343: contig of 1878 bp in length  
45443: gap of unknown length  
45987: contig of 1544 bp in length  
47087: gap of unknown length  
49404: contig of 2317 bp in length  
49504: gap of unknown length  
50823: contig of 1319 bp in length



AAGLRBPVPVPATADPELLAESDEAKVORAAAGVGGSLRAQVERLRVLEORERRRBBE					
ORDSFEGERLAWQAEKEQVIRYQKQHQNYIQMYRRNRQLEQELQQLSLEARELAD					
LGLAEQAPCICLEBITATEI"					
BASE COUNT	355 a	619 c	668 g	293 t	
ORIGIN					
Query Match	17.5%	Score 314;	DB 9;	Length 1935;	
Best Local Similarity	59.6%;	Pred. No. 3.7e-35;			
Matches 617;	Conservative 0;	Mismatches 365;	Indels 54;	Gaps 3;	
Qy	796	CTGAGCAGCAGAGCTGTGTTGAGAGGAGGCGCCCTCCAGAAAGCTGCAGCGCAGCTTTGAG	855		
Db	910	CTGAGAGAAAGCTCCGAGACCGGAGCAGCAGCTTTCAGCAGCTGCGGACAGTCTGGAC	969		
Qy	856	GAGAAGGAGCTTGCTCCAGCTCGCCTTACGAGAGCGCGCGCGCTGCGAGGACGAG	915		
Db	970	GAGAAATGAGGCTACCATGTGCGCAGCATACGAGAGCGGCGCAGCGCACTGCGCAGGAG	1029		
Qy	916	CTGAGAGGCCCGGAGCCCAAGAGCGGCNAACAGCTCAAGCAGGCTCGCAAGAGCGCAG	975		
Db	1030	CGTGAGGCCCTGCGAGAGGACTGTGCGGCCCGCAGGCACAGCGGGC-----ACAG	1077		
Qy	976	CGCGCGCAGCAGCTCTGCACCTGCAGTCTCTGAGCTTCAGCAGGAGAAAGCGCAGCTC	1035		
Db	1078	CGGCCCCACAGCTGTCTGAGCTGAGTGTTCAGCTGCGAGGAGGAGCGGCAATTG	1137		
Qy	1036	CGGAGGAGCTCGAGAGCCTCATGAAGAGCAGGACCTCTGCGAGACCAAGCTCAGGTCC	1095		
Db	1138	CAGGACGACTTCGACAGCTGCTGCAGAGCGCGNACAGCTGGAGCGCGCTGCGCCACC	1197		
Qy	1096	TACGAGAGGAGAGACACAGCTTCGCGCCCGCTGAGAGAGCCAGTGGAGGTGTC	1155		
Db	1198	TTGAGCGCGGAGCAGCGGAGCTCGGCGCAGGCTTGAGAGAGACCAAGTGGAGGTGTC	1257		
Qy	1156	CAGAAGTCAGCGCAGATCTCCCTCTGAGCAGCAGCTGAGAGGTCCAGCAGCGAGTG	1215		
Db	1258	CAGAAATCAGCGCAGATCTCCCTGCTGAAGCAGCAGCTGAAAGAGTCTCAGCGCAGAGCTG	1317		
Qy	1216	AACGCCAAGGCTAGCAGATCTGGGTCTCAAGGCACAGCTGAAGGACACCGCGGGCAAG	1275		
Db	1318	GTGCAGAAGGCGCAGAGCTGTGTGCTCTGCGGGTGGCGCTCGGAGGCCGCTGCTACG	1377		
Qy	1276	CTGAGAGGCTGAGCTGAGACCCAGACCTGAGAGGCGCCCTGCGCACCAAGGGCTG	1335		
Db	1378	CTGCGGCTCAGTAGGCGCGTGTGCGCGGGTCTACAGGAGCGCGCGAGCTCGGAGCTG	1437		
Qy	1336	GAGCTGAGGTCTGTGAGATGAGCTGAGCGCAGAGAACAGAGCGGAGCTCTGCGG	1395		
Db	1438	GAGCTGGAAGCTGTTCAGGAGCTGAGCGCACCGCCAGGAGCTGAGCAGCTGCGG	1497		
Qy	1396	GAGAAGGTGAACCTGCTGAGCAGGAGCTGAGCAGGCTGCGGGC-----	1439		
Db	1498	GAGAAGCTGGCAGTTGGATGCTGAGCGCGCGGACTCCGAGAGCCCTGTGCCACCT	1557		
Qy	1440	-----CCAGGCGCCCTTGCGCGCGCAGATGGGCGCCGCCACCTTC	1479		
Db	1558	GCCACCGCTGACCCATCTCTCGCAGAGAGTGATGAGGCAAGTGACGCGGCGAGCA	1617		
Qy	1480	CCGAGGACGCTCCCTCGCTCAGCGGAGCTGAGCGGCTGCGGCGGAGCTCGGGAG	1539		
Db	1618	GCGGGGTTGGGGGAGCTTTCGCGGCCAGGTGAGGCGATTTGCGGGTGGAGCTCGACGG	1677		
Qy	1540	GAGCGGCAAGCCATGACCAAGATGCTCTCGGGCTTCAGACATGAGCGGCTCGTGTGAAG	1599		
Db	1678	GAGCGCGCGGGGTGAGGAGCAGCGGACAGCTTTGAGGGGGAGCGGCTGGCTTGGCAG	1737		
Qy	1600	GAGGAGAGGAGAGAGGTGATTAGTACAGAAACAGCTGACAGAGCTACGTGGCCATG	1659		
Db	1738	GCAGAGAAGGAGCAGGTGATCCGCTACCAAGACAGCTGACAGCAACAATACATCCAGATG	1797		
Qy	1660	TACCAGCGGACACGCGCTGAGAGAGCCCTGACAGCTGGCACCTGGGGACAGCGCC	1719		
Db	1798	TACCGGCGCAACCGGACAGTAGACAGAGCTGACAGAGCTCAGCCTGGAGCTGGAGGCC	1857		

Qy	1720	GGGAGAGCCCTGGAGGTTGACCTGG-----AAGGGGCTGACATCCCTCAGAGGACATC	1773			
Db	1858	CGGAGAGCTCGCTGACCTGGGCTGGCGAGCAGGCCCCCTGCATCTGCTCTGGAGGATC	1917			
Qy	1774	ATAGCCACTGAGATCT	1789			
Db	1918	ACTGCTACTGAGATCT	1933			
RESULT 13						
LOCUS	HS0805394	2275 bp	mRNA	linear	PRI 12-JUL-2002	
DEFINITION	Homo sapiens mRNA; cDNA DKFZp761K1711 (from clone DKFZp761K1711).					
ACCESSION	AL834338					
VERSION	AL834338.1 GI:21739969					
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 2275)					
TITLE	Ansorge,W., Winkner,U., Mewes,H.W., Weil,B. and Wiemann,S.					
JOURNAL	Direct Submission					
COMMENT	Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY					
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.					
	This clone (DKFZp761K1711) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.					
FEATURES	Location/Qualifiers					
source	1..2275					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/chromosome="10"					
	/clone="DKFZp761K1711"					
	/tissue_type="amygdala"					
	/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"					
	/dev_stage="adult"					
gene	1..2275					
	/gene="DKFZp761K1711"					
CDS	<1..1597					
	/gene="DKFZp761K1711"					
	/note="similarity to FEZ1 (Homo sapiens), different splicing pattern"					
	/codon_start=2					
	/product="hypothetical protein"					
	/protein_id="CAD39005.1"					
	/db_xref="GI:21739970"					
	/translation="EKILIRPTAFKVLPPKPRGAPSLPFMPGRATGLSGSGSLTOLFGPSSSSSSSSAADKPLAFSGWASGCPSTLSDGRNLSLSLPTVSTGAEPTTSPGGLHPHSGRGALPGKRGVTPGSHSDGRSSSKSTGLSGRVAGGLGSGTRASPDSGSGERSFPFPPPPDEALLHCVLGKLRDREAEKLRDLSLDNEATMCQAYEERHWRHREAREADCAQAQQQLQVQLQFQQKQESQAEIVQKGSGLVREGLRRCATLREORELGPRLTEKWEVCQKSGEISLLKQKESQAEIVQKGSGLVALVAVREARATLVSEGRARGLQEAARELEACSOELRHROBAEOLREKAGOLDAAAGLRPVPVPATADPELLAESDEAKVORAAAGVGGSLRAQVERLRVLEORERRGEQRDSFEGERLAWQAEKEQVIRYQKQHQNYIQMYRRNRQLEQELQQLSLEARELADLGLAEQAPCICLEBITATEI"					
	2236..2241					
polyA_signal	/gene="DKFZp761K1711"					
	2257					
polyA_site	/gene="DKFZp761K1711"					
BASE COUNT	455 a 697 c 739 g 384 t					
ORIGIN						
Query Match	17.5%; Score 314; DB 9; Length 2275;					

Best Local Similarity 59.6%; Pred. No. 3.6e-35;		Matches 617; Conservative 0; Mismatches 365; Indels 54; Gaps 3;	
Qy	796	CTGAGCAGAGCTGTTGGAGAGGAGGCGCCCTCCAGAGAGCTGCACGACGAGCTTTGAG	855
Db	572	CTGGAAGAAAGCTCCGAGACCGGAGGAGCTTTCAGAGCTGCGGACAGTCTGGAC	631
Qy	856	GAGAAGAGCTTGCTTCAGCTGCTCCAGCTGCTACAGAGAGCGGCGCGCTGCGAGGAGCGAG	915
Db	632	GAGAAATGAGGCTACCATGTGTCAGGACATACAGAGAGCGGACGCGCTGCGAGCGAGAG	691
Qy	916	CTGAGGCGCGGAGCCCAAGAGCGGCAACAAAGCTCAAGCAGGAGCTCGCAGAGAGCCAG	975
Db	692	CTGAGGCGCTGCGAGGAGCTGTCGCGCCAGGACAGCGGC-----ACAG	739
Qy	976	CGCGCGCAGCAGGCTTCAGCTGAGTATGTCAGCTTTCAGAGGAGAGCGGACGCTC	1035
Db	740	CGGCGCCCAACAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCGAGGAGAGCGGCAATTG	799
Qy	1036	CGGAGGAGCTCGAGGCTCATGAGGAGCAGGAGCTGCTGAGGAGCCAGCTCAGGCTCAGGCTC	1095
Db	800	CAGGACGACTTGCACAGCTGCTGAGGAGCGGCAACAGCTGAGGAGCGGCTGCGCCACC	859
Qy	1096	TACGAGAGGAGAGAACCCAGCTTTCGCGCCCGCTGAGGAGAGCCAGTGGAGGTGTGC	1155
Db	860	TTGAGCGGAGCAGCGGAGCTGCGCGGAGGCTTGGAGAGCAAGTGGAGGTGTGC	919
Qy	1156	CAGAAGTCAGCGGAGATCTCCTCTGAAGCAGCAGCTGAAGAGTGTCCAGACCGAGGTG	1215
Db	920	CAGAAATCAGCGGAGATCTCCTCTGTAAGCAGCAGCTGAAAGAGTCTCAGGACAGAGCTG	979
Qy	1216	AACGCCAAGCTAGGAGATCTGCTGCTTCAAGCAGCAGCTGAAGGACAGCGCGGGCAAG	1275
Db	980	GTGAGAGGAGCGAGCAGCTGCTGCTGCTGCGGCTGCGGAGGCGGCTGCTGCTGCTGCTG	1039
Qy	1276	CTGAGGCGCTGAGCTGAGGACCCAGGACCTGAGGCGGCGCTGCGGACCAAGGCGCTG	1335
Db	1040	CTGCGGCTCAGTGGAGGCGCTGCGGCGGCTTACAGGAGCGCGCGGAGCTGCGGAGCTG	1099
Qy	1336	GAGTGGAGGCTGTGAGAAATGAGCTGAGCGGCAAGAGAACGAGGCGGAGCTGCTGCGG	1395
Db	1100	GAGCTGGAAGCTGTTCAGGAGCTGCGGAGCTGAGGACACCGCAGGAGAGTGAAGCTGCGG	1159
Qy	1396	GAGAGGTGACCTGCTGAGCAGGAGCTGAGGAGCTGCGGCGGCTGCGGCGGCTGCGGCGG	1439
Db	1160	GAGAAAGCTGGGAGTGGATGCTGAGGCGGCGGAGCTGCGGAGCGGCGGCTGCTGCTGCTG	1219
Qy	1440	-----CCAGCGCGCGCTGCGCGCGGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1479
Db	1220	CCACCGCTGACCAATCTCTCTGCGAGAGTGTAGGCGCAAGTGCAGCGGCGGAGCA	1279
Qy	1480	CCGAGGAGCTGCTGCTGCGGAGCTGAGGAGCTGCGGCGGCTGCGGCGGAGCTGCGGAG	1539
Db	1280	GCGGCGGCTGGGCGGAGCTGCGGCGGCGGAGCTGAGGAGCTGCGGCTGCGGAGCTGCGG	1339
Qy	1540	GAGCGGAGCCATGACAGATGCTGCGGCTTCCAGCATGAGCGGCTGCTGCTGGAAG	1599
Db	1340	GAGCGCGCGGCTGAGGAGCAGCGGAGCTTGGAGGCGGAGCGGCTGCTGCTGCGGAG	1399
Qy	1600	GAGGAGAGGAGAGGCTGATTTCAGTACCAAGAAACAGCTGCGAGCAGAGCTAGCTGCGCATG	1659
Db	1400	CGAGAGAGGAGCAGCTGATCGCTACCAAGAGCAGCTGCGAGCAACTATCATCAGATG	1459
Qy	1660	TACAGCGGAGACCGAGCTGAGAGCGCTGAGCAGCTGCGAGCTGCGGAGCGGAGCGGCGG	1719
Db	1460	TACCGCGCAACCGGAGCTGAGCAGGAGCTGCGAGAGCTGCGAGCTGCGGAGCTGCGGAG	1519
Qy	1720	GGGAGCGCTTGGAGTTGACTGG-----AAGGGCTGACATCCCTTACGAGGAGATC	1773
Db	1520	CGGAGCTGCTGAGCTGCTGCGGCTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1579
Qy	1774	ATAGCACTGAGATCT 1789	

Db	1580	ACTGCTACTGAGATCT 1595
RESULT 14		
BC006212		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
REMARK		
COMMENT		

BC006212 2766 bp mRNA linear PRI 12-JUL-2001  
 Homo sapiens, clone MGC:2586 IMAGE:3161855, mRNA, complete cds.  
 BC006212.1 GI:13623228  
 MGC.  
 Homo sapiens.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2766)  
 Strausberg, R.  
 Direct Submission  
 Submitted (09-APR-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantropop, S., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,  
 Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 5 Row: k Column: 1  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis, GenomeScan gene prediction.

FEATURES	Location/Qualifiers
source	1..2766
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="MGC:2586 IMAGE:3161855"
	/tissue_type="Brain, neuroblastoma"
	/clone_lib="NIH MGC 19"
	/lab_host="DH10B-R"
	/note="Vector: pOTB7"
	1138..2088
	/codon_start=1
	/product="Unknown (protein for MGC:2586)"
	/protein_id="AAH06212.1"
	/db_xref="GI:13623229"
	/translation="MCQAYEERHWRERAREALDCAAQRAQRAQQLQLQVFL QOKRQLODPAOLLOEROLERRCATLERQELGPRLETKWEVCCKGEISLLKQ QKESQNELVQKSELVALRVALREARATLRVSEGRARGLQEARARELEACSOEL QRHQEABQUREKAGQDLAAAGRPFPVPTADRLPFAEASDEAKVQRAAGVGGSL RAQVERLRVLEORRERQDSFEGERLAWQAEQVIRYQKQLOHNYIQMYRRNR ALBQELQLQSLLEARELADLGLAEAPCICLEITATEI"

BASE COUNT 545 a 873 c 883 g 465 t

ORIGIN

Query Match 17.5%; Score 314; DB 9; Length 2766;  
 Best Local Similarity 59.6%; Pred. No. 3.5e-35;  
 Matches 617; Conservative 0; Mismatches 365; Indels 54; Gaps 3;

```
Qy 796 CTGAGCAGAAAGCTGTTGGAGGAGGAGCGCGCCCTCCAGAAAGCTGCAGCGCAGCTTTGAG 855
Db 1063 CTGGAAGGAAAGCTCCGAGACCGGAGCAGAGCTTACAGCAGCTGCGGACAGTCTGGAC 1122
Qy 856 GAGAGGAGCTTGCTCCAGCTCGCCTACAGGAGCGCGCGCTGCAGGACGAG 915
Db 1123 GAGATGAGCTTACCATGTGTCAGGATACAGGAGCGCGGACACTGGCAGGAG 1182
Qy 916 CTGAGGCGCGGAGCCCAAGCGGCAACAAAGCTCAAGCAGCGGCTTCGCAAGAGCGCAG 975
Db 1183 CGTGAGGCGCTCGAGAGAGTGTGCGCGCCAGGCACACGCGGC-----ACAG 1230
Qy 976 CGCGCGAGAGCTCTCCACCTGACAGTACTGAGCTTACAGGAGAGAGCGGAGCTC 1035
Db 1231 CGGCGCCCAACAGCTGCTGAGCTGAGGCTTCCAGCTGCGAGCAGGAGGCGCAATTG 1290
Qy 1036 CGGAGGAGCTCGAGAGCTCATGAAGCAGCAGGACCTGCTGGAGACCAAGCTCAGGTCC 1095
Db 1291 CAGGAGCTTGCACAGCTGCTGAGGAGCGGACAGCTGGAGCGGCTGCGCCACC 1350
Qy 1096 TACGAGGAGGAGAACACAGCTTCGCGCCCGCTGGAGGAGACCCAGTGGAGGTGTC 1155
Db 1351 TTGAGCGGAGCAGCGGAGCTCGGCGCGAGCTTGAAGGAGACCAAGTGGAGGTGTC 1410
Qy 1156 CAGAGTCAAGCGGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAGTCCAGCAGGAGTG 1215
Db 1411 CAGAAATCAGCGAGATCTCCCTGCTGAAGCAGCAGCTGAAAGAGTCTCAGCGCAGAGCTG 1470
Qy 1216 AACGCCAAGGCTACGAGATCTTGGTCTCAAGCAGCAGCTGAAGGACACCGCGGGCAG 1275
Db 1471 GTGCAGAAAGGCGAGCTGGTGTCTGCGGCTGCGGCTGCGGAGCGGCGCTGCTACG 1530
Qy 1276 CTGAGGCGCTTGGAGCTCAGGACCCAGGACCTGGAGGCGGCTCGCGCACCAAGGCGCTG 1335
Db 1531 CTGCGGCTCAGTAGGCGGCTGCGGCGGTCTACAGGAGCGCGCGAGCTCGGGAGCTG 1590
Qy 1336 GAGCTGAGGCTGTGAGAAATGAGCTGAGCGGAGCTGAGCGGCTGCGGCGGAGCTGTCGG 1395
Db 1591 GAGCTGGAAGCTGTCTCCAGGAGCTGCGGAGCAGCAGCAGGAGCTGAGCAGCTGCGG 1650
Qy 1396 GAGAGGTGACCTCTGAGCAGGAGCTGAGGAGCTGCGGCG----- 1439
Db 1651 GAGAAAGTGGGAGTTGGATGCTGAGGCGCGGAGCTCGGAGCGCCCTGTGCCACCT 1710
Qy 1440 -----CCAGCGCCCTGCGCGCGGACATGGGCGCGCCACCTTC 1479
Db 1711 GCCACCGTGACCATTCCTCTGCGAGAGTGATGAGGCGCAAGTGACGCGGCGAGCA 1770
Qy 1480 CCGAGGAGCTCCCTGCGGAGCTGAGCGGCTGCGGCGGAGCTGCGGAG 1539
Db 1771 GCGGCGGTTGGGCGGAGCTTGGGCGGCGGAGCTGAGGCGGATTTGCGGCTGAGCTGCGG 1830
Qy 1540 GAGCGGCAAGCCATGACAGATGCTCGGCGCTTCCAGCATGAGCGCTGCTGGAG 1599
Db 1831 GAGCGCGCGGCGGTGAGGAGCGGAGCTTTGAGGCGGAGCGGCTGCGCTGGCAG 1890
Qy 1600 GAGGAGAGGAGAGGATGATTCAGTACCAAGAAACAGCTGCGAGCAGAGCTAGCTGGCCATG 1659
Db 1891 CGAGAGAGGAGAGGATGATCGGTACCAAGAGAGCTGCGAGCAACTATCATCCAGATG 1950
Qy 1660 TACGAGCGGAACCAAGCGCTCGAGAGCGCTGCGAGCAGCTGGCAGCTGGGAGCAGCGCC 1719
Db 1951 TACCGGCGCAACCGCAGCTAGAGCAGGAGCTGAGCAGAGCTCAGCTTGGAGCTGGAGGCC 2010
Qy 1720 GGGAGCGCTTGGAGGTTGACTGG-----AAGGGGCTGACATCCCTACGAGCAATC 1773
Db 2011 CGGAGCTCGTGTGACCTGGGCGCTGCGCGAGCAGGCGCCCTGCTCATCTGCTGGAGAGATC 2070
Qy 1774 ATAGCCACTGAGATCT 1789
Db 2071 ACTGCTACTGAGATCT 2086
```

```
RESULT 15
LOCUS AB011124
DEFINITION Homo sapiens mRNA for KIAA0552 protein, complete cds.
ACCESSION AB011124
VERSION AB011124.1 GI:3043627
KEYWORDS KIAA0552 protein.
SOURCE Homo sapiens male brain cDNA to mRNA, clone lib:pBluescriptII SK
plus clone:HH0869.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro
JOURNAL DNA Res. 5 (1), 31-39 (1998)
MEDLINE 98290545
REFERENCE 2 (bases 1 to 5257)
AUTHORS Ohara,O., Nagase,T. and Ishikawa,K.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:cdmainfo@kazusa.or.jp, Tel: +81-438-52-3913,
Fax: +81-438-52-3914)
FEATURES
Location/Qualifiers
source 1..5257
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HH0869"
/sex="male"
/tissue type="brain"
/clone_lib="pBluescriptII SK plus"
gene 1..5257
/gene="KIAA0552"
CDS 1409..3430
/gene="KIAA0552"
/codon_start=1
/product="KIAA0552 protein"
/protein_id="BAA25478.1"
/db_xref="GI:3043628"
/translaton="MAKLETPVRADPDRPDLAFAPRPSLGPDPRLAMGSGV
AHQEPFAMKVGTRTGGGSGGSPGPRGSGARRPGRYPSEDKGLANSLVNGE
LRGSDHDTVCGNVGSSGSSSGGSKAPPOVREPSHPKLLATSKLDQCSPLVR
PSAFKVPVKNFHSQMNLCPQNTGTEGCGGLGGLDKSTMTWTPAGSGGSLD
SGRSLTSLTSSYSOHLAPUSATSHINRITAGSYGSGSGSGSGSGYQDGTG
DSGRASKSGSSSMRPHLGSGGGGLPFAACSPSPSALIQLERLWKEQE
VAALRRLEQSEAQAQVLEERQKAWERELAEFGCGSKLQVRAQRAQGLQLQ
VLRLQDKQLQBEAARLMRQBELEDKVAQKEADFLPRIETKWEVCQKAGEIS
LLKQOLKSDADYSOKLSEIVGLRSOLREGRASIREKEOLLSLRDSFSSKQASLELG
EGELPACLPALTTPVDPAEPDQALATCESDEAKMRRAQGVAAASLVSVDGEAEG
ESGTRALRRVRLQAEALAAARRARERQGSFAEERRVWLEKEKVLEYQKQLQLSYV
EMTORNQLRRLRRERGAAGGATPTTPHQHEKKAWTPPSRLRIESTE1"
BASE COUNT 997 a 1618 c 1673 g 969 t
ORIGIN
Query Match 17.5%; Score 314; DB 9; Length 5257;
Best Local Similarity 58.2%; Pred. No. 3,1e-35;
Matches 619; Conservative 0; Mismatches 400; Indels 45; Gaps 2;
Qy 710 GCAGCTGGGCCACTCGCAACAGGACAGAGGCGCCCTCGTGTGTCGCGCTCCCGCATCT 769
Db 2292 GCCACCTGGGCTCTGGGAGGGCGGAGGTGAGCGCTTTCGCGGCTTCCTACCGC 2351
Qy 770 CCACGAGCGAGTGCGAGCATCCAGGAGCTGGAGCAGAGAGCTGTTGGAGAGGAGCGGCC 829
Db 2352 CCTCCCCAGTGCACTCATCCAGGAGCTGGAGGAGCGGCTGTGGAGAGGAGGAGGAGG 2411
Qy 830 TCCAGAGCTGCGAGCGAGCTTTGAGGAGAGGAGCTTGCCTCCAGCTGCGCTACGAGG 889
Db 2412 TGGCAGCTCTGCGGCGGAGCGCTGAGCAGAGCGAGCGGCGGTGTGGCCCGAGGTACTG 2471
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 22:38:41 ; Search time 412.345 Seconds  
(without alignments)  
9781.434 Million cell updates/sec

Title: US-09-513-888C-3  
Perfect score: 1791  
Sequence: 1 atggcgagcgtcagtagcct.....tcatagccactgagatctga 1791

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1791	100.0	1791	21	AAA64509
2	1791	100.0	5492	21	AAA64508
3	1643	91.7	1722	21	AAA64515
4	1581.4	88.3	1692	21	AAA64514
5	1427	79.7	1614	21	AAA64512
6	1372.4	76.6	1512	21	AAA64513
7	808.8	45.2	2411	22	AAS30637
8	808.8	45.2	2411	22	AAS28699
9	807.2	45.1	9048	22	AAA64507
					CDNA sequence enco
					CDNA sequence of c
					Nucleotide sequenc
					Nucleotide sequenc
					Nucleotide sequenc
					DNA encoding novel
					Genomic sequence #
					Nucleotide sequenc

10	562.6	31.4	633	21	AAA64511	Nucleotide sequenc
11	560	31.3	560	22	AAS30595	DNA encoding novel
12	560	31.3	560	22	AAS28127	Novel cDNA encodin
13	269.8	15.1	1623	24	ABO54361	Human ovarian anti
14	258.2	14.4	1617	21	AAF15625	Human prostate can
15	249	13.9	404	21	AAA64510	Nucleotide sequenc
16	148.4	8.3	657	22	AAF22599	Human breast cance
17	146	8.2	1926	21	AAA50254	Epstein Barr virus
18	146	8.2	1926	22	AAF82902	EBV tethering prot
19	146	8.2	2580	21	AAA75454	Nucleotide sequenc
20	146	8.2	2580	24	AAI64275	Epstein-Barr virus
c 21	146	8.2	5452	20	AAZ30923	Anti-sense strand
c 22	146	8.2	8705	20	AAZ23778	Vector pshuttle DN
23	146	8.2	9600	19	AAV21683	Vector plasmid pCM
24	146	8.2	10380	20	AAZ22248	Nucleotide sequenc
25	146	8.2	10596	14	AAQ51731	Plasmid pCisEBON f
26	146	8.2	10596	17	AAAT40348	Plasmid pCisEBON f
27	146	8.2	10596	20	AAI15650	Nucleotide sequenc
c 28	146	8.2	16080	21	AAAS9553	DNA clone pCEK Cl.
29	144.2	8.1	799	19	AAV55831	Nucleotide sequenc
30	138.4	7.7	3489	21	AAA30290	Kaposi's sarcoma-a
31	138.4	7.7	3489	22	AAF82901	Nucleotide sequenc
32	138.4	7.7	3489	24	ABA93487	Kaposi's sarcoma-a
c 33	138.4	7.7	32207	20	AAV73805	KSHV LUR DNA (nucl
c 34	138.4	7.7	137507	19	AAV19941	KSHV long unique c
35	134	7.5	1925	20	AAZ90924	Epstein Barr Virus
36	131.4	7.3	795	19	AAV55830	FLGA insert stabil
37	131.2	7.3	2108	24	ABL67774	Oesophagus cancer
38	125.2	7.0	9551	20	AAZ22301	CDNA encoding a hu
39	118.2	6.6	2744	16	AAQ98470	MISPI-containing p
40	118	6.6	1588	22	AAH26304	Spider silk protei
41	117.4	6.6	2004	18	AAAT85356	Nephila clavipes s
42	117	6.5	2338	12	AAQ14183	N.clavipes draglin
c 43	114	6.4	561	24	ABK62287	Rat sequence diffe
44	113.8	6.4	2338	19	AAV23249	Nephila clavipes s
45	113.8	6.4	2338	21	AAZ38195	N. clavipes spider

ALIGNMENTS

RESULT 1	
AAA64509	
ID	AAA64509 standard; cDNA; 1791 BP.
AC	AAA64509;
XX	
DT	02-JAN-2001 (first entry)
XX	
DE	CDNA sequence encoding a human FEZ1 polypeptide.
XX	
KW	Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW	tumour proliferation; tubulin; microtubule; protein BFL gamma;
KW	tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW	cell growth; cell shape; cell rigidity; cell motility; DNA replication;
XX	tumorigenesis; tumour survival; metastasis; ss.
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1791
FT	/*tag= a
FT	/product= "FEZ1"
XX	
PN	WO2000050565-A2.
XX	
PD	31-AUG-2000.
XX	
PF	25-FEB-2000; 2000WO-US04950.
XX	
PR	25-FEB-1999; 99US-0121537.
XX	
PA	(UYJE-) UNIV JEFFERSON THOMAS.

```
XX Croce CM, Ishii H;
XX PI
XX WPI; 2000-558396/51.
DR P-PSDB; AAB08715.
XX
XX New polynucleotide homologous with a portion of one strand of the human
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as
PT cancer -
XX
XX Claim 7; Fig 5I; 255pp; English.
XX
XX The present sequence encodes a human FEZ1 polypeptide. FEZ1 is a
CC tumour suppressor gene, located at chromosome location 8p22. Decreased
CC or no expression of FEZ1 is detected in a variety of cancer cells.
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC also interacts with tubulin, with microtubules, and with protein
CC E1-gamma. Post-translational phosphorylation and dephosphorylation
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC expression are useful for inducing cells to proliferate. Compounds
CC which modulate FEZ1 association with tubulin are useful for alleviating
CC tubulin hyper- or hypo- polymerisation disorders, such as those
CC associated with aberrant initiation of mitosis, modulation of the
CC initiation and rate of cell proliferation and cell growth, modulation of
CC cell shape, cell rigidity, cell motility, rate and stage of cellular
CC DNA replication, intracellular distribution of organelles, metastatic
CC potential of cell and cellular transformation from a non-cancerous to
CC cancerous phenotype. Compounds which modulate FEZ1 binding and
CC phosphorylation are also useful for alleviating a disorder, such as
CC tumorigenesis, tumour survival, growth and metastasis.
XX
SQ Sequence 1791 BP; 393 A; 561 C; 591 G; 246 T; 0 other;

Query Match 100.0%; Score 1791; DB 21; Length 1791;
Best Local Similarity 100.0%; Pred. No. 7.4e-288;
Matches 1791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGGGCAGCGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60
DB |||||||
DB 1 ATGGGCAGCGTCAGTAGCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60
OY 61 TCGCAGTACAAGCTCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTTCGACGGG 120
DB |||||||
DB 61 TCGCAGTACAAGCTCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTTCGACGGG 120
OY 121 CTGCTGAGGTTGGCTTCTCCAGAGACTCCGGTCAAGGAGTCCAGCTCCAAATGGGC 180
DB |||||||
DB 121 CTGCTGAGGTTGGCTTCTCCAGAGACTCCGGTCAAGGAGTCCAGCTCCAAATGGGC 180
OY 181 AAGACGGAAGACTTCTTACATCAAGTCAAGCAAGAGCCGGGCTCCCATCACCCA 240
DB |||||||
DB 181 AAGACGGAAGACTTCTTACATCAAGTCAAGCAAGAGCCGGGCTCCCATCACCCA 240
OY 241 GATTACAGGCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 300
DB |||||||
DB 241 GATTACAGGCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCG 300
OY 301 TCCACACCCCAAGCTCATGCCCTTCTCCATAGCTAGAAATGGGCTCCGAGAGGGT 360
DB |||||||
DB 301 TCCACACCCCAAGCTCATGCCCTTCTCCATAGCTAGAAATGGGCTCCGAGAGGGT 360
OY 361 GCAGTGAGGCCACAGCTTCAAGCTGTGTGCCACAGCTCAGGAGCCATCTGCACATCC 420
DB |||||||
DB 361 GCAGTGAGGCCACAGCTTCAAGCTGTGTGCCACAGCTCAGGAGCCATCTGCACATCC 420
OY 421 TCCCGGAGAGTCCAGCCACAGCTGCAACCCCGCCCTTCCACAGCAAGCCCAAGAGCAG 480
DB |||||||
DB 421 TCCCGGAGAGTCCAGCCACAGCTGCAACCCCGCCCTTCCACAGCAAGCCCAAGAGCAG 480
OY 481 GAGCTGAGGCTGGGCTGTGTCTGGGGGCTGTGAGACTCCGGCCGGAATCTCCATGTCC 540
DB |||||||
DB 481 GAGCTGAGGCTGGGCTGTGTCTGGGGGCTGTGAGACTCCGGCCGGAATCTCCATGTCC 540

541 AGCTGCCACACACAGCACCAGCAGCAGCTACAGCTGGACCCCGCTGGTCACACCCGCTG 600
DB |||||||
541 AGCTGCCACACACAGCACCAGCAGCAGCTACAGCTGGACCCCGCTGGTCACACCCGCTG 600
OY 601 GGACCCCAAGCCGTTTGGGGGCTCCGCCACAACATACCCAGGGGATCGTCTCTCAG 660
DB |||||||
601 GGACCCCAAGCCGTTTGGGGGCTCCGCCACAACATACCCAGGGGATCGTCTCTCAG 660
OY 661 GACAGCAACATGATGAGCCTGAGGCTCTGCTCTTCTCCAGCAGGAGTACGAGCTGGC 720
DB |||||||
661 GACAGCAACATGATGAGCCTGAGGCTCTGCTCTTCTCCAGCAGGAGTACGAGCTGGC 720
OY 721 CACTCGAACAGGAGCAGCAAGGGCCCTCGTGTGTCCGCTCCCGCATCTCCACGAGCAG 780
DB |||||||
721 CACTCGAACAGGAGCAGCAAGGGCCCTCGTGTGTCCGCTCCCGCATCTCCACGAGCAG 780
OY 781 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTGTGAGAGGAGGGCGCCCTCCAGAAAGCTG 840
DB |||||||
781 TGCAGCATCCAGGAGCTGGAGCAGAAAGCTGTGTGAGAGGAGGGCGCCCTCCAGAAAGCTG 840
OY 841 CAGCGCAGCTTGGAGAGAGGAGCTTGCCTCCAGCCTGCGCTACGAGGAGGAGCGCCGCG 900
DB |||||||
841 CAGCGCAGCTTGGAGAGAGGAGCTTGCCTCCAGCCTGCGCTACGAGGAGGAGCGCCGCG 900
OY 901 CGCTGCAGGAGCAGAGCTGGAGGGGCCCGAGCCCAAGCGCGCAACAAAGCTCAAGCAGGCC 960
DB |||||||
901 CGCTGCAGGAGCAGAGCTGGAGGGGCCCGAGCCCAAGCGCGCAACAAAGCTCAAGCAGGCC 960
OY 961 TCGCAGAAAGCAGCAGCGCGCAGAGCTGTGCACTGAGGTACTGAGCTTACAGCAG 1020
DB |||||||
961 TCGCAGAAAGCAGCAGCGCGCAGAGCTGTGCACTGAGGTACTGAGCTTACAGCAG 1020
OY 1021 GAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGGAGGAGCTCTCTGAG 1080
DB |||||||
1021 GAGAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGGAGGAGCTCTCTGAG 1080
OY 1081 ACCAAGCTCAGGTCCTACGAGAGGAGAGAACACAGCTTCCGCCCCGGCTCGAGGAGACC 1140
DB |||||||
1081 ACCAAGCTCAGGTCCTACGAGAGGAGAGAACACAGCTTCCGCCCCGGCTCGAGGAGACC 1140
OY 1141 CAGTGGAGGTGTGCCAAGAGTCAGGCGAGATCTCTCTCTTCTGAAAGCAGCAGCTGAAAG 1200
DB |||||||
1141 CAGTGGAGGTGTGCCAAGAGTCAGGCGAGATCTCTCTCTTCTGAAAGCAGCAGCTGAAAG 1200
OY 1201 TCCAGAGCGAGGTGAACGCCCAAGCTTAGCGAGATCTGGGTCTCAGGGCACAGCTGAAG 1260
DB |||||||
1201 TCCAGAGCGAGGTGAACGCCCAAGCTTAGCGAGATCTGGGTCTCAGGGCACAGCTGAAG 1260
OY 1261 GACACGGGGGCAAGCTGGAGGGCTTGAGCTGAGGAGCCCAAGGAGCTGGAGGGCGCCCTG 1320
DB |||||||
1261 GACACGGGGGCAAGCTGGAGGGCTTGAGCTGAGGAGCCCAAGGAGCTGGAGGGCGCCCTG 1320
OY 1321 CGCAACCAAGGGCTTGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGGCGCAAGAAACGAG 1380
DB |||||||
1321 CGCAACCAAGGGCTTGGAGCTGGAGGTCTGTGAGAAATGAGCTGCGGCGCAAGAAACGAG 1380
OY 1381 GCGGAGCTGTGCGGGGAGAGGTGAACCTGCTGAGCAGAGGAGCTGAGGAGCTGCGGGCC 1440
DB |||||||
1381 GCGGAGCTGTGCGGGGAGAGGTGAACCTGCTGAGCAGAGGAGCTGAGGAGCTGCGGGCC 1440
OY 1441 CAGCGCCGCTTGGGCCCGCAGCATGGGGCGCCACCTTCCCGGAGGAGCTGCTGCGCTG 1500
DB |||||||
1441 CAGCGCCGCTTGGGCCCGCAGCATGGGGCGCCACCTTCCCGGAGGAGCTGCTGCGCTG 1500
OY 1501 CAGCGGAGCTGAGAGCGGCTGCGGGCGAGCTGCGGGAGGAGCGGCAAGGCATGACGAG 1560
DB |||||||
1501 CAGCGGAGCTGAGAGCGGCTGCGGGCGAGCTGCGGGAGGAGCGGCAAGGCATGACGAG 1560
OY 1561 ATGTCTCTGGGCTTCCAGCATGAGCGGCTGTGTGAGAGGAGGAGAGGAGAGGTGATT 1620
DB |||||||
1561 ATGTCTCTGGGCTTCCAGCATGAGCGGCTGTGTGAGAGGAGGAGAGGAGAGGTGATT 1620
OY 1621 CAGTACCAGAAACAGCTGCGAGAGCTACGTGGCCATGTACCGAGCGGAAACGAGCGCTG 1680
```

Db 1621 CAGTACCAAGAACAGCTCAGCAGAGCTACGTGGCCATGTATCCAGCGAAACAGCGCCTG 1680  
Qy 1681 GAGAAAGGCGCTGACGAGCTGGCAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGAC 1740  
Db 1681 GAGAAAGGCGCTGACGAGCTGGCAGCTGGGACAGCGCGGGAGCCCTTGGAGGTTGAC 1740  
Qy 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGGCACTGAGATCTGA 1791  
Db 1741 CTGGAAGGGGCTGACATCCCTTACGAGGACATCATAGGCACTGAGATCTGA 1791

RESULT 2

AAA64508  
ID AAA64508 standard; cDNA; 5492 BP.  
AC  
XX  
AC  
XX  
AA64508;

DT 02-JAN-2001 (first entry)

DE cDNA sequence of the wild type human FEZ1 gene.

XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein Efi-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX Homo sapiens.

OS WO2000050565-A2.

XX 31-AUG-2000.

PN 25-FEB-2000; 2000WO-US04950.

XX 25-FEB-1999; 99US-0121537.

PR (UVEJ-) UNIV JEFFERSON THOMAS.

XX Croce CM, Iehli H;

XX WPI; 2000-558396/51.

XX New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -

PS Example 2; Fig 5B; 255pp; English.

XX The present sequence represents the cDNA sequence of the human FEZ1 gene.  
CC FEZ1 is a tumour suppressor gene, located at chromosome location 8p22.  
CC Decreased or no expression of FEZ1 is detected in a variety of cancer  
CC cells. Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC Efi-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.

XX Sequence 5492 BP; 1137 A; 1704 C; 1565 G; 1086 T; 0 other;

Query Match 100.0%; Score 1791; DB 21; Length 5492;

Best Local Similarity 100.0%; Pred. No. 7.8e-288;

Matches 1791; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGGCGAGCGTCAGTAGCTCTCATCTCCGGGCGACAGCTTCCACAGCAAGCACTCCCGGGCT 60  
Db 112 ATGGGCGAGCGTCAGTAGCTCTCATCTCCGGGCGACAGCTTCCACAGCAAGCACTCCCGGGCT 171  
Qy 61 TCGCAGTAAAGCTGCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTTCCACGGG 120  
Db 172 TCGCAGTAAAGCTGCGCAAGTCTCTCCACCTCAAGAAAGCTCAACCGGTATTTCCACGGG 231  
Qy 121 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGC 180  
Db 232 CTGCTGAGGTTTGGCTTCTCCAGGACTCCGGTCAAGCAAGTCCAGCTCCAAAATGGGC 291  
Qy 181 AAGAGCGAAGACTTCTTCTACATCAAGGTCAAGAAAGCTCCGGGCTCCCATCACCCA 240  
Db 292 AAGAGCGAAGACTTCTTCTACATCAAGGTCAAGAAAGCTCCGGGCTCCCATCACCCA 351  
Qy 241 GATTACAGGCACTGTCTCCAGCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCG 300  
Db 352 GATTACAGGCACTGTCTCCAGCGGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCG 411  
Qy 301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGGCTCCGAGAGGGT 360  
Db 412 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGGCTCCGAGAGGGT 471  
Qy 361 GCAGTGAAGGCGCCACAGCGCTTCAAGCTGTGTCACAGGTCAGGAGCCATCTGCACTCC 420  
Db 472 GCAGTGAAGGCGCCACAGCGCTTCAAGCTGTGTCACAGGTCAGGAGCCATCTGCACTCC 531  
Qy 421 TCCCGGAGAGTGCCAGCCACAGCTGCACCCCCCTCCAGACAAGCCCAAGAGGAG 480  
Db 532 TCCCGGAGAGTGCCAGCCACAGCTGCACCCCCCTCCAGACAAGCCCAAGAGGAG 591  
Qy 481 GAGCTGAAGCGCTGGCTGTGCTCTGGGGCGCTGTCAAGCTCCGGCCGGAATCCATGTCC 540  
Db 592 GAGCTGAAGCGCTGGCTGTGCTCTGGGGCGCTGTCAAGCTCCGGCCGGAATCCATGTCC 651  
Qy 541 AGCCTGCCCCACACAGCAGCAGCAGCAGCTACAGCTGGAGCCCGCTGGTCAACACCCGTG 600  
Db 652 AGCCTGCCCCACACAGCAGCAGCAGCAGCTACAGCTGGAGCCCGCTGGTCAACACCCGTG 711  
Qy 601 GGACCCCAAGCCGTTTGGGGGCTCCGCCCAACAACATCACCCAGGGGATCGTCTCTCCAG 660  
Db 712 GGACCCCAAGCCGTTTGGGGGCTCCGCCCAACAACATCACCCAGGGGATCGTCTCTCCAG 771  
Qy 661 GACAGCAACATGATGAGCCTGAAGCTGTGCTTCTCCGAGGAGGCTAGCAAGCTGGGC 720  
Db 772 GACAGCAACATGATGAGCCTGAAGGCTGTGCTTCTCCGAGGAGGCTAGCAAGCTGGGC 831  
Qy 721 CACTCGAACAAGGCGAGCAAGGGGCCCTCGTGTGTGCTCGCTCCCCCATCTCCACGGACGAG 780  
Db 832 CACTCGAACAAGGCGAGCAAGGGGCCCTCGTGTGTGCTCGCTCCCCCATCTCCACGGACGAG 891  
Qy 781 TGCAGCATCCAGGAGCTGGAGCAGAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTG 840  
Db 892 TGCAGCATCCAGGAGCTGGAGCAGAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAAGCTG 951  
Qy 841 CAGCGAGCTTTGAGGAGAGGAGCTTGCCTCCAGCTTGGCTTACGAGGAGGGCGCCGG 900  
Db 952 CAGCGAGCTTTGAGGAGAGGAGCTTGCCTCCAGCTTGGCTTACGAGGAGGGCGCCGG 1011  
Qy 901 CGCTGCAAGGACAGCTGGAGGGGCCCGAGGCCCAAGAGCGGCAACAAGCTCAAGCAGAGCC 960  
Db 1012 CGCTGCAAGGACAGCTGGAGGGGCCCGAGGCCCAAGAGCGGCAACAAGCTCAAGCAGAGCC 1071  
Qy 961 TCGCAGAAAGCGCAGCGCGCAGCAGGTCTCTGCACTTGCAGGTACTGCAAGCTTCCAGCAG 1020  
Db 1072 TCGCAGAAAGCGCAGCGCGCAGCAGGTCTCTGCACTTGCAGGTACTGCAAGCTTCCAGCAG 1131  
Qy 1021 GAGAAAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGAGCTTGTGTGAG 1080  
Db 1132 GAGAAAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGAGCTTGTGTGAG 1191

QY 1081 ACCAAGCTCAGGTCTTACGAGAGGAGAAAGACCAGCTTCGGCCCGCGCTCGAGAGACC 1140  
Db 1192 ACCAAGCTCAGGTCTTACGAGAGGAGAGAACAGCTTCGGCCCGCGCTCGAGAGAGACC 1251  
QY 1141 CAGTGGAGGTGTGCGAGAGTCAAGGAGATCTCCCTCTTGAAGCAGCAGCTGAAGGAG 1200  
Db 1252 CAGTGGAGGTGTGCGAGAGTCAAGGAGATCTCCCTCTTGAAGCAGCAGCTGAAGGAG 1311  
QY 1201 TCCACAGCGAGGTGAACCCAGAGCTAGCAGATCTCTGGTCTCAAGGACAGCTGAAG 1260  
Db 1312 TCCACAGCGAGGTGAACCCAGAGCTAGCAGATCTCTGGTCTCAAGGACAGCTGAAG 1371  
QY 1261 GACACGGCGGCGAGCTGGAGGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTG 1320  
Db 1372 GACACGGCGGCGAGCTGGAGGCTGAGCTGAGGACCCAGGACCTGGAGGCGCCCTG 1431  
QY 1321 CGACCAAGGCGCTGAGCTGAGGTCTGTGAGATGAGCTGACGCGCAAGAAAGCAG 1380  
Db 1432 CGACCAAGGCGCTGAGCTGAGGTCTGTGAGATGAGCTGACGCGCAAGAAAGCAG 1491  
QY 1381 GCGAGCTGTGCGGAGAAAGTGAACCTGTGAGAGCAGAGCTGACAGAGCTGCGGACC 1440  
Db 1492 GCGAGCTGTGCGGAGAAAGTGAACCTGTGAGAGCAGAGCTGACAGAGCTGCGGACC 1551  
QY 1441 CAGCGCGGCTGGCGCGGACATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCCCTG 1500  
Db 1552 CAGCGCGGCTGGCGCGGACATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCCCTG 1611  
QY 1501 CAGCGGAGCTGGAGCGCTGCGGCGGAGCTGCGGAGGAGCGCAAGGCGCATGACCG 1560  
Db 1612 CAGCGGAGCTGGAGCGCTGCGGCGGAGCTGCGGAGGAGCGCAAGGCGCATGACCG 1671  
QY 1561 ATGCTCTCGGCTTCCAGCATGAGCGCTCGTGTGAAGAGGAGAAAGAGTGTATT 1620  
Db 1672 ATGCTCTCGGCTTCCAGCATGAGCGCTCGTGTGAAGAGGAGAAAGAGTGTATT 1731  
QY 1621 CAGTACCAAGAACAGCTCAGCAGAGCTACGTGGCCATGTACCAGCGAAACCGGCGCTG 1680  
Db 1732 CAGTACCAAGAACAGCTCAGCAGAGCTACGTGGCCATGTACCAGCGAAACCGGCGCTG 1791  
QY 1681 GAGAAGGCGCTGACAGCTGCGAGCTGGGACAGCGCGGGGAGCCCTTGGAGTTGAC 1740  
Db 1792 GAGAAGGCGCTGACAGCTGCGAGCTGGGACAGCGCGGGGAGCCCTTGGAGTTGAC 1851  
QY 1741 CTGGAAGGGGCTGACATCCCTTACGAGACATCATAGCCACTGAGATCTGA 1791  
Db 1852 CTGGAAGGGGCTGACATCCCTTACGAGACATCATAGCCACTGAGATCTGA 1902

RESULT 3  
AAA64515  
ID AAA64515 standard; cDNA; 1722 BP.  
XX  
AC AAA64515;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Nucleotide sequence of truncated FEZ1 transcript G3612.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1722  
FT /tag= a  
FT /product= "truncated FEZ1"

PN WO2000050565-A2.  
XX 31-AUG-2000.  
XX 25-FEB-2000; 2000WO-US04950.  
XX 25-FEB-1999; 99US-0121537.  
XX (UYJE-) UNIV JEFFERSON THOMAS.  
FA Croce CM, Ishii H;  
XX WPI: 2000-558396/51.  
XX P-PSDB; AAB08722.  
DR New polynucleotide homologous with a portion of one strand of the human  
XX FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
XX cancer -  
XX Disclosure; Fig 5H; 255pp; English.  
XX  
CC The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 1722 BP; 381 A; 541 C; 563 G; 237 T; 0 other;  
Query Match 91.78; Score 1643; DB 21; Length 1722;  
Best Local Similarity 96.1%; Pred. No. 2.3e-263;  
Matches 1722; Conservative 0; Mismatches 0; Indels 69; Gaps 1;  
QY 1 ATGGGCGAGCGTCAGTAGCCTCATCTCGGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60  
Db 1 ATGGGCGAGCGTCAGTAGCCTCATCTCGGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60  
QY 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120  
Db 61 TCGCAGTACAGCTGCGCAAGTCTCCACCTCAAGAGCTCAACCGGTATTCGACGGG 120  
QY 121 CTGCTGAGGTTGGCTTCTCCAGAGCTCCGGTCAACGCGAGTCCAGTCCAAATGGC 180  
Db 121 CTGCTGAGGTTGGCTTCTCCAGAGTCCGGTCAACGCGAGTCCAGTCCAAATGGC 180  
QY 181 AAGACGCAAGACTTCTTCTACATCAAGGTCAAGCAAGAGCCCGGGGCTCCCATCACC 240  
Db 181 AAGACGCAAGACTTCTTCTACATCAAGGTCAAGCAAGAGCCCGGGGCTCCCATCACC 240  
QY 241 GATTACAGGCACTGTCTCAGCGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCG 300  
Db 241 GATTACAGGCACTGTCTCAGCGGGATTTAGGGGCGCAGGCTGGGGTGGACTTTGACCCG 300  
QY 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGT 360  
Db 301 TCCACACCCCGCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGT 360  
QY 361 GCAGTGAGGGCCACAGGCTTCAAGCCTGTGTGTCGACGGTCAAGGAGCCATCTGCACTCC 420

Db 361 GAGGTGAGGCGCACAGCCTTCAAGCCTGTGTGTGCCACCGGTGAGGAGCCATCTGCACTCC 420  
Qy 421 TCCCGGAGAGTGCAGCAGCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAG 480  
Db 421 TCCCGGAGAGTGCAGCAGCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGAGCAG 480  
Qy 481 GAGCTGAAGCCTGGCCTGTGTCTGTGGGCGGTGTGAGATCTCGGCGGAACTCCATGTCC 540  
Db 481 GAGCTGAAGCCTGGCCTGTGTCTGTGGGCGGTGTGAGATCTCGGCGGAACTCCATGTCC 540  
Qy 541 AGCTGCCCCACACAGCAGCACAGCTGACAGCTGACAGCTGGACCCGCTGTCACACCCGCTG 600  
Db 541 AGCTGCCCCACACAGCAGCACAGCTGACAGCTGACAGCTGGACCCGCTGTCACACCCGCTG 600  
Qy 601 GGACCCACAAGCGTTTTGGGGGCTCCGCCACACATCACCAGGGGATCGTCTCCAG 660  
Db 601 GGACCCACAAGCGTTTTGGGGGCTCCGCCACACATCACCAGGGGATCGTCTCCAG 660  
Qy 661 GACAGCAACATGATGAGCCTGAGGCTCTGTCTTCTCCGAGGAGGTAGCAAGCTGGGC 720  
Db 661 GACAGCAACATGATGAGCCTGAGGCTCTGTCTTCTCCGAGGAGGTAGCAAGCTGGGC 720  
Qy 721 CACTCGACAGGAGCTGGAGCAGAACTGTGTGTGTCGCTCCCTCCATCTCCACGACGAG 780  
Db 721 CACTCGACAGGAGCTGGAGCAGAACTGTGTGTGTCGCTCCCTCCATCTCCACGACGAG 780  
Qy 781 TGCAGATCCAGGAGCTGGAGCAGAACTGTGTGAGAGGGGCGCCCTCCAGAACTG 840  
Db 781 TGCAGATCCAGGAGCTGGAGCAGAACTGTGTGAGAGGGGCGCCCTCCAGAACTG 840  
Qy 841 CAGCGAGCTTTGAGGAGAGAGCTTGCCTCCAGCCTGGCTACGAGGAGCGCGCGG 900  
Db 841 CAGCGAGCTTTGAGGAGAGAGCTTGCCTCCAGCCTGGCTACGAGGAGCGCGCGG 900  
Qy 901 CGCTGCAGGAGAGCTGGAGGCGCGGAGCCCAAGCGGCAACAGCTCAAGCAGGCC 960  
Db 901 CGCTGCAGGAGAGCTGGAGGCGCGGAGCCCAAGCGGCAACAGCTCAAGCAGGCC 960  
Qy 961 TCGCAGAGAGAGCTGAGCGCGCAGCAGCTCTGACCTGACAGTACTGACAGTTCAGCAG 1020  
Db 961 TCGCAGAGAGAGCTGAGCGCGCAGCAGCTCTGACCTGACAGTACTGACAGTTCAGCAG 1020  
Qy 1021 GAGAAAGCGGAGCTCCGCGAGAGCTCGAGAGCTCATGAAGAGCAGGAGCTGCTGAG 1080  
Db 1021 GAGAAAGCGGAGCTCCGCGAGAGCTCGAGAGCTCATGAAGAGCAGGAGCTGCTGAG 1080  
Qy 1081 ACCAGCTCAGTCTTACAGAGGAGAGACAGCAGCTTCGGCCCGCGCTGGAGAGACC 1140  
Db 1081 ACCAGCTCAGTCTTACAGAGGAGAGACAGCAGCTTCGGCCCGCGCTGGAGAGACC 1140  
Qy 1141 CAGTGGGAGGTGTGCCAGAGTCAAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG 1200  
Db 1141 CAGTGGGAGGTGTGCCAGAGTCAAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG 1200  
Qy 1201 TCCAGAGCGAGGTGAACCAAGGCTAGCAGATCTTGGGTCTCAAGCAGCAGCTGAAG 1260  
Db 1201 TCCAGAGCGAGGTGAACCAAGGCTAGCAGATCTTGGGTCTCAAGCAGCAGCTGAAG 1260  
Qy 1261 GACACCGGGGCAAGCTGGAGGCGCTGAGAGTGAAGACCCAGGAGCTGGAGGCGCCCTG 1320  
Db 1261 GACACCGGGGCAAGCTGGAGGCGCTGAGAGTGAAGACCCAGGAGCTGGAGGCGCCCTG 1320  
Qy 1321 CGCACAGGCGCTGAGGCTGTGAGATGTGAGATGAGCTGACGCGCAAGAAAGCAG 1380  
Db 1321 CGCACAGGCGCTGAGGCTGTGAGATGTGAGATGAGCTGACGCGCAAGAAAGCAG 1380  
Qy 1381 GCGAGAGCTGCTGGGAGAGAGTGAACCTGTGGAGCAGGAGCTGAGGAGCTGGGCGCC 1440  
Db 1381 GCGAGAGCTGCTGGGAGAGAGTGAACCTGTGGAGCAGGAGCTGAGGAGCTGGGCGCC 1440  
Qy 1441 CAGCGCGCTGCGCCCGCAGATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCGCCTG 1500  
Db 1441 CAGCGCGCTGCGCCCGCAGATGGGCGCGCCACCTTCCCGAGGAGCTCCCTGCGCCTG 1500

Qy 1501 CAGCGGAGCTGGAGCGGCTCGGCGCGAGCTGCGGGAGGAGCGCAAGGCCATGACCCAG 1560  
Db 1501 CAGCGGAGCTG----- 1512  
Qy 1561 ATGTCTCGGGCTTCCAGCATGAGCGGCTCTGTGGAAAGGAGGAGAAAGAGTGTATT 1620  
Db 1513 -----GAGCGGCTCTGTGTGGAAGGAGGAGAAAGAGTGTATT 1551  
Qy 1621 CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAAGCGGAAACAGCGCCTG 1680  
Db 1552 CAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAAGCGGAAACAGCGCCTG 1611  
Qy 1681 GAGAAGGCCCTGCAGCAGCTGGCGGACAGCGCGCGGAGGCCCTTTGGAGGTTGAC 1740  
Db 1612 GAGAAGGCCCTGCAGCAGCTGGCGGACAGCGCGCGGAGGCCCTTTGGAGGTTGAC 1671  
Qy 1741 CTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1791  
Db 1672 CTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACTGAGATCTGA 1722

RESULT 4  
AAH64514  
ID AAA64514 standard; cDNA; 1692 BP.  
XX AC AAA64514;  
XX AC  
DT 02-JAN-2001 (first entry)  
XX  
DE Nucleotide sequence of truncated FEZ1 transcript G3611.  
XX Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein EPI-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1692  
FT /\*tag= a  
FT /product= "truncated FEZ1"  
XX  
PN WO200050565-A2.  
XX  
PD 31-AUG-2000.  
XX  
PF 25-FEB-2000; 2000WO-US04950.  
XX  
PR 25-FEB-1999; 99US-0121537.  
XX  
PA (UYJE-) UNIV JEFFERSON THOMAS.  
XX  
PI Croce CM, Ishii H;  
XX WPI; 2000-558396/51.  
DR P-PSDB; AAB08721.  
XX  
PT New polynucleotide homologous with a portion of one strand of the human  
FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
cancer -  
XX  
PS Disclosure; Fig 5G; 255pp; English.  
XX  
CC The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC EPI-gamma. Post-translational phosphorylation and dephosphorylation

modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are useful for inducing cells to proliferate. Compounds which modulate FEZ1 association with tubulin are useful for alleviating tubulin hyper- or hypo- polymerisation disorders, such as those associated with aberrant initiation of mitosis, modulation of the initiation and rate of cell proliferation and cell growth, modulation of cell shape, cell rigidity, cell motility, rate and stage of cellular DNA replication, intracellular distribution of organelles, metastatic potential of cell and cellular transformation from a non-cancerous to cancerous phenotype. Compounds which modulate FEZ1 binding and phosphorylation are also useful for alleviating a disorder, such as tumorigenesis, tumour survival, growth and metastasis.

SQ Sequence 1692 BP; 380 A; 523 C; 553 G; 236 T; 0 other;

Query Match 88.3%; Score 1581.4; DB 21; Length 1692;  
 Best Local Similarity 94.4%; Pred. No. 3.7e-253;  
 Matches 1691; Conservative 0; Mismatches 1; Indels 99; Gaps 1;

```

QY 1 ATGGCAGGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60
DB 1 ATGGCAGGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60
QY 61 TCGCAGTACAAGTCGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTTCCGACGGG 120
DB 61 TCGCAGTACAAGTCGCGAAGTCTCCACCTCAAGAGCTCAACCGGTATTTCCGACGGG 120
QY 121 CTGCTGAGTTGGTTCTTCCAGGACTCCGGTCAACGCAAGTCCAGTCCCAAAATGGGC 180
DB 121 CTGCTGAGTTGGTTCTTCCAGGACTCCGGTCAACGCAAGTCCAGTCCCAAAATGGGC 180
QY 181 AAGCGGAAGACTTCTTCTACATCAAGTCAAGCAAGAGCCCGGGCTCCCATCACCCA 240
DB 181 AAGCGGAAGACTTCTTCTACATCAAGTCAAGCAAGAGCCCGGGCTCCCATCACCCA 240
QY 241 GATTACAGGCACTCTCAGCGGGATTTAGGGGCCAGGCTGGGGTGGATTTTGACCCG 300
DB 241 GATTACAGGCACTCTCAGCGGGATTTAGGGGCCAGGCTGGGGTGGATTTTGACCCG 300
QY 301 TCCACACCCCAAGCTCATGCGCTTCTCAATCAGTCTAGAAATGGGCTCCGAGAGGGT 360
DB 301 TCCACACCCCAAGCTCATGCGCTTCTCAATCAGTCTAGAAATGGGCTCCGAGAGGGT 360
QY 361 GCAGTGAAGCCACAGCTTCAAGCTTGCTGCCACCGTCAAGGAGCCATCTGCACCTCC 420
DB 361 GCAGTGAAGCCACAGCTTCAAGCTTGCTGCCACCGTCAAGGAGCCATCTGCACCTCC 420
QY 421 TCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGCAAGCCCAAGGAGCAG 480
DB 421 TCCCGGAGAGTCCAGCCACAGCTGCACCCCGCCCTCCAGCAAGCCCAAGGAGCAG 480
QY 481 GAGCTGAAGCTGGCTGTGCTCTGGGGCGTGTGAGACTCCGCGCGGAACTCCATGTCC 540
DB 481 GAGCTGAAGCTGGCTGTGCTCTGGGGCGTGTGAGACTCCGCGCGGAACTCCATGTCC 540
QY 541 AGCCTGCCACACAGCAGCAGCAGCTACAGCTACGAGTGGAGCCCGCTGGTACACCCGTG 600
DB 541 AGCCTGCCACACAGCAGCAGCAGCTACAGCTACGAGTGGAGCCCGCTGGTACACCCGTG 600
QY 601 GGACCCACAAAGCGTTTGGGGGCTCCGCCCCAACACATCACCCAGGATCGTCTCCAG 660
DB 601 GGACCCACAAAGCGTTTGGGGGCTCCGCCCCAACACATCACCCAGGATCGTCTCCAG 660
QY 661 GACAGCAATGATGAGCTGAAGGCTGTGCTTCTCCGACGAGGCTAGCAAGCTGGGC 720
DB 661 GACAGCAATGATGAGCTGAAGGCTGTGCTTCTCCGACGAGGCTAGCAAGCTGGGC 720
QY 721 CACTCGAACAGCGACAGAGGCCCTCGTGTGTGCTGCCCTCCCTCCACGACGAG 780
DB 721 CACTCGAACAGCGACAGAGGCCCTCGTGTGTGCTGCCCTCCCTCCACGACGAG 780
QY 781 TGCAGCATCCAGGAGCTGGAGCAGAGCTGTGTGGAGAGGAGGGCGCCCTCCAGAGCTG 840

```

```

DB 781 TGCAGCATCCAGGAGCTGGAGCAGAGAGCTGTGTGGAGAGGAGGGCGCCCTCCAGAAGCTG 840
QY 841 CAGCGCAGCTTTGAGGAGAGAGGAGCTTGCCTCCAGCTTGCCTTACGAGGAGCGCCCGG 900
DB 841 CAGCGCAGCTTTGAGGAGAGAGGAGCTTGCCTCCAGCTTGCCTTACGAGGAGCGCCCGG 900
QY 901 CGCTGCAAGCAGCAGCTGAGAGGCGCGGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCC 960
DB 901 CGCTGCAAGCAGCAGCTGAGAGGCGCGGAGCCCAAGGCGGCAACAAGCTCAAGCAGGCC 960
QY 961 TCGCAGAAAGCAGCAGCTGAGAGGCGCGGAGCTTCTGCACCTGAGGTTACTGAGCTTCAAG 1020
DB 961 TCGCAGAAAGCAGCAGCTGAGAGGCGCGGAGCTTCTGCACCTGAGGTTACTGAGCTTCAAG 1020
QY 1021 GAGAAAGCAGCTTCCGCGAGGAGCTTCAGAGCTTCAAGAGGAGGAGGAGCTTCTGAG 1080
DB 1021 GAGAAAGCAGCTTCCGCGAGGAGCTTCAGAGCTTCAAGAGGAGGAGGAGCTTCTGAG 1080
QY 1081 ACCAAGCTCAGGCTTCTACGAGAGGAGGAGAGCAAGCTTTCGGCCCGCGCTGGAGGAG 1140
DB 1081 ACCAAGCTCAGGCTTCTACGAGAGGAGGAGAGCAAGCTTTCGGCCCGCGCTGGAGGAG 1140
QY 1141 CAGTGGAGGTTGTGCCAAGAGTCAAGGAGATCTTCCCTCTGAAAGCAGCAGCTGAAGGAG 1200
DB 1141 CAGTGGAGGTTGTGCCAAGAGTCAAGGAGATCTTCCCTCTGAAAGCAGCAGCTGAAGGAG 1200
QY 1201 TCCAGACGAGGTTGAACGCGCAAGGCTTAGCGAGATCTTGGTCTCAAGGAGCAGCTGAAG 1260
DB 1201 TCCAGACGAGGTTGAACGCGCAAGGCTTAGCGAGATCTTGGTCTCAAGGAGCAGCTGAAG 1260
QY 1261 GACACGCGGGCAGCTTGGAGGGCTTGGAGCTGAGGAGCTTGGAGGGCGCCCTG 1320
DB 1261 GACACGCGGGCAGCTTGGAGGGCTTGGAGCTGAGGAGCTTGGAGGGCGCCCTG 1320
QY 1321 CGCACCAAGGCTTGGAGCTTGTGAGAAATGAGCTGACGCGCAAGAAAGCAG 1380
DB 1321 CGCACCAAGGCTTGGAGCTTGTGAGAAATGAGCTGACGCGCAAGAAAGCAG 1380
QY 1381 GCGAGCTGTGCGGGAGAGGTTGAACCTGCTGAGAGCAGGAGCTGAGGAGCTGGGGCC 1440
DB 1381 GCGAGCTGTGCGGGAGAGGTTGAACCTGCTGAGAGCAGGAGCTGAGGAGCTGGGGCC 1440
QY 1441 CAGGCGCCCTGGCGCGCAGCATGGGGCGCCACCTTCCCGGAGGAGCTCCCTGCCCTG 1500
DB 1441 CAGGCGCCCTGGCGCGCAGCATGGGGCGCCACCTTCCCGGAGGAGCTCCCTGCCCTG 1500
QY 1501 CAGCGGAGCTGGAGCGGCTGCGGGCGAGCTGCGGAGGAGCGGCAAGGCAATGACAG 1560
DB 1501 CAGCGGAGCTGGAGCGGCTGCGGGCGAGCTGCGGAGGAGCGGCAAGGCAATGACAG 1560
QY 1561 ATGTCTCGGGCTTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGGAGGAGGAGGAGT 1620
DB 1561 ATGTCTCGGGCTTCCAGCATGAGCGGCTGCTGTGGAAGGAGGAGGAGGAGGAGGAGT 1620
QY 1621 CAGTACCAAGAAACAGCTGACAGCAGCTTACGTGGCCATGTACCAAGGAGGAGGAGGAG 1680
DB 1621 CAGTACCAAGAAACAGCTGACAGCAGCTTACGTGGCCATGTACCAAGGAGGAGGAGGAG 1680
QY 1681 GAGAAAGGCTTGCAGCAGCTGCGGAGGAGCGCGGGAGCGCCCTTCGAGGTTGAC 1740
DB 1681 GAGAAAGGCTTGCAGCAGCTGCGGAGGAGCGCGGGAGCGCGCCCTTCGAGGTTGAC 1740
QY 1741 CTGGAAGGGCTGACATCCCTTACGAGGAGCAGTACGAGGAGGAGGAGGAGGAGGAGT 1791
DB 1741 CTGGAAGGGCTGACATCCCTTACGAGGAGCAGTACGAGGAGGAGGAGGAGGAGGAGT 1791
QY 1642 CTGGAAGGGGCTGATCCTCCCTTACGAGGAGCAGTACGAGGAGGAGGAGGAGGAGT 1692
DB 1642 CTGGAAGGGGCTGATCCTCCCTTACGAGGAGCAGTACGAGGAGGAGGAGGAGGAGT 1692

```

RESULT 5  
 AAA64512  
 ID AAA64512 standard; cDNA; 1614 BP.  
 XX  
 AC  
 AA64512;  
 XX

DT	02-JAN-2001 - (first entry)
XX	Nucleotide sequence of truncated FEZ1 transcript T8D145M4.
DE	
XX	Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW	tumour proliferation; tubulin; microtubule; protein EF1-gamma;
KW	tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW	cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW	tumorigenesis; tumour survival; metastasis; ss.
XX	
OS	Homo sapiens.
XX	
PH	Key Location/Qualifiers
FT	1..633
CDS	/tag= a
FT	/product= "truncated FEZ1"
FT	
PN	WO200050565-A2.
XX	
PD	31-AUG-2000.
XX	
PF	25-FEB-2000; 2000WO-US04950.
XX	
PR	25-FEB-1999; 99US-0121537.
XX	
PA	(UYJE-) UNIV JEFFERSON THOMAS.
XX	
PI	Croce CM, Ishii H;
XX	
DR	WPI: 2000-558396/51.
XX	
PT	P-PSDB; AAB08719.
XX	
PS	New polynucleotide homologous with a portion of one strand of the human
XX	FEZ1 gene, useful for alleviating abnormal cell proliferation such as
XX	cancer -
XX	Disclosure; Fig 5E; 255pp; English.
XX	
CC	The present sequence encodes a truncated human FEZ1 polypeptide. The
CC	encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour
CC	suppressor gene, located at chromosome location 8p22. Decreased
CC	or no expression of FEZ1 is detected in a variety of cancer cells.
CC	Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1
CC	also interacts with tubulin, with microtubules, and with protein
CC	EF1-gamma. Post-translational phosphorylation and dephosphorylation
CC	modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene
CC	expression are useful for inducing cells to proliferate. Compounds
CC	which modulate FEZ1 association with tubulin are useful for alleviating
CC	tubulin hyper- or hypo- polymerisation disorders, such as those
CC	associated with aberrant initiation of mitosis, modulation of the
CC	initiation and rate of cell proliferation and cell growth, modulation of
CC	cell shape, cell rigidity, cell motility, rate and stage of cellular
CC	DNA replication, intracellular distribution of organelles, metastatic
CC	potential of cell and cellular transformation from a non-cancerous to
CC	cancerous phenotype. Compounds which modulate FEZ1 binding and
CC	phosphorylation are also useful for alleviating a disorder, such as
CC	tumorigenesis, tumour survival, growth and metastasis.
XX	
SQ	Sequence 1614 BP; 367 A; 500 C; 522 G; 225 T; 0 other;
	Query Match 79.7%; Score 1427; DB 21; Length 1614;
	Best Local Similarity 90.1%; Pred. No. 1.3e-227;
	Matches 1614; Conservative 0; Mismatches 0; Indels 177; Gaps 1;
QY	1 ATGGGCAGCGTCAGTAGCCTCATCTCCGSCCACACAGCTTCCACACAAGCACTGCCGGCT 60
DB	 1 ATGGGCAGCGTCAGTAGCCTCATCTCCGSCCACACAGCTTCCACACAAGCACTGCCGGCT 60
QY	61 TCGCAGTACAAGCTCGGAAGTCTCCACCCTCAAGAAGCTCAACCGGTATTTCCGACGGG 120
DB	 61 TCGCAGTACAAGCTCGGAAGTCTCCACCCTCAAGAAGCTCAACCGGTATTTCCGACGGG 120
QY	121 CTGCTGAGGTTTGCTTCTTCCAGAGACTCCCGTCAACGGCAAGTCCAGCTCCAAATGGGC 180
DB	 121 CTGCTGAGGTTTGCTTCTTCCAGAGACTCCCGTCAACGGCAAGTCCAGCTCCAAATGGGC 180



Db 1201 TCCAGCGGAGGTGAACCCAGGCTAGCGAGATCTGGGTCTCAAGGCACACGCTGAAG 1260  
Qy 1261 GACACGGGGGCAAGCTGGAGGGCTTGAGCTGAGACCCAGGACCTTGAGGGGGCCCTG 1320  
Db 1261 GACACGGGGGCAAGCTGGAGGGCTTGAGCTGAGGACCCAGGACCTTGAGGGGGCCCTG 1320  
Qy 1321 CGCACCAAGGGCTGGAGGCTTGAGGCTTGAGGATGAGCTGACGCGCAAGAAAGAG 1380  
Db 1321 CGCACCAAGGGCTGGAGGCTTGAGGCTTGAGGATGAGCTGACGCGCAAGAAAGAG 1380  
Qy 1381 GCGGAGCTGCTGCGGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCC 1440  
Db 1381 GCGGAGCTGCTGCGGGGAG----- 1399  
Qy 1441 CAGGCCGCCCTGGCCCCCGACATCGGGGCCGCCACCTTCCCGAGGAGCTCCCTGCCCCCTG 1500  
Db 1400 ----- 1399  
Qy 1501 CAGCGGGAGCTGGAGCGGCTCGGGGCCGAGCTCGGGGAGGAGCGCAAGGCCATGACCAG 1560  
Db 1400 ----- 1399  
Qy 1561 ATGCTCTCGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGGTGATT 1620  
Db 1400 -----AGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGGTGATT 1443  
Qy 1621 CAGTACCAAGAAAGCTGCAGCAGAGCTACCTGGCCATGTACCGAGGAAACCGGCGCTG 1680  
Db 1444 CAGTACCAAGAAAGCTGCAGCAGAGCTACCTGGCCATGTACCGAGGAAACCGGCGCTG 1503  
Qy 1681 GAGAGGCCCTGCGAGCAGCTGCGCAGCTGGGCGACAGCGCGGGGAGCCCTTGAGGTTGAC 1740  
Db 1504 GAGAGGCCCTGCGAGCAGCTGCGCAGCTGGGCGACAGCGCGGGGAGCCCTTGAGGTTGAC 1563  
Qy 1741 CTGAAGGGGCTGACATCCCTACGAGCAGCATATAGCCACTGAGATCTGA 1791  
Db 1564 CTGAGGGGCTGACATCCCTACGAGCAGCATATAGCCACTGAGATCTGA 1614

RESULT 6  
ID AAA64513  
XX AAA64513 standard; cDNA; 1512 BP.  
AC AAA64513;  
XX  
DT 02-JAN-2001 (first entry)  
XX  
DE Nucleotide sequence of truncated FEZ1 transcript D14.  
XX  
KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;  
KW tumour proliferation; tubulin; microtubule; protein Efr1-gamma;  
KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;  
KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;  
KW tumorigenesis; tumour survival; metastasis; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 1..1512  
FT CDS /\*tag= a  
FT /product= "truncated FEZ1"  
XX  
PN MO200050565-A2.  
XX  
XX 31-AUG-2000.  
XX  
XX 25-FEB-2000; 2000WO-US04950.  
XX  
XX 25-FEB-1999; 99US-0121537.  
XX (U9JE-) UNIV JEFFERSON THOMAS.  
XX  
XX Croce CW, Ishii H;

XX WPI; 2000-558396/51.  
DR P-PSDB; AAB08720.  
XX  
PT New polynucleotide homologous with a portion of one strand of the human  
PT FEZ1 gene, useful for alleviating abnormal cell proliferation such as  
PT cancer -  
XX  
PS Disclosure; Fig 5F; 255pp; English.  
XX  
CC The present sequence encodes a truncated human FEZ1 polypeptide. The  
CC encoding mRNA is transcribed by tumour cells. FEZ1 is a tumour  
CC suppressor gene, located at chromosome location 8p22. Decreased  
CC or no expression of FEZ1 is detected in a variety of cancer cells.  
CC Expression of FEZ1 inhibits tumour growth and proliferation. FEZ1  
CC also interacts with tubulin, with microtubules, and with protein  
CC Efr1-gamma. Post-translational phosphorylation and dephosphorylation  
CC modulates the effect of the FEZ1 protein. Inhibitors of FEZ1 gene  
CC expression are useful for inducing cells to proliferate. Compounds  
CC which modulate FEZ1 association with tubulin are useful for alleviating  
CC tubulin hyper- or hypo- polymerisation disorders, such as those  
CC associated with aberrant initiation of mitosis, modulation of the  
CC initiation and rate of cell proliferation and cell growth, modulation of  
CC cell shape, cell rigidity, cell motility, rate and stage of cellular  
CC DNA replication, intracellular distribution of organelles, metastatic  
CC potential of cell and cellular transformation from a non-cancerous to  
CC cancerous phenotype. Compounds which modulate FEZ1 binding and  
CC phosphorylation are also useful for alleviating a disorder, such as  
CC tumorigenesis, tumour survival, growth and metastasis.  
XX  
SQ Sequence 1512 BP; 335 A; 483 C; 481 G; 213 T; 0 other;  
Query Match 76.6%; Score 1372.4; DB 21; Length 1512;  
Best Local Similarity 95.4%; Pred. No. 1.5e-218;  
Matches 1425; Conservative 0; Mismatches 66; Indels 3; Gaps 1;  
Qy 1 ATGGGAGCGTCAGTAGCCTCATCTCGGGCCACAGCTTCCACAGCAAGCAGCTCGCGGCT 60  
Db 1 ATGGGAGCGTCAGTAGCCTCATCTCGGGCCACAGCTTCCACAGCAAGCAGCTCGCGGCT 60  
Qy 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120  
Db 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGGG 120  
Qy 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTACGCGAAGTCCAGTCCAAATGGGC 180  
Db 121 CTGCTGAGGTTGGCTTCTCCAGGACTCCGGTACGCGAAGTCCAGTCCAAATGGGC 180  
Qy 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240  
Db 181 AAGAGCGAAGACTTCTTCTACATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240  
Qy 241 GATTACAGGCACTGTCCAGCGGGATTAGGGGGCCAGGCTGGGGTGGATTTGACCCG 300  
Db 241 GATTACAGGCACTGTCCAGCGGGATTAGGGGGCCAGGCTGGGGTGGATTTGACCCG 300  
Qy 301 TCCACACCCCCCAAGCTCATGCCCCCTTCCAAATCAGCTAGAAAATGGGCTCCGAGAAGGT 360  
Db 301 TCCACACCCCCCAAGCTCATGCCCCCTTCCAAATCAGCTAGAAAATGGGCTCCGAGAAGGT 360  
Qy 361 CGAGTAGGGCCACAGAGCTTCAAGCCTGTGTGTCACGGTCCAGGAGGAGGCTCCGACTCC 420  
Db 361 CGAGTAGGGCCACAGAGCTTCAAGCCTGTGTGTCACGGTCCAGGAGGAGGCTCCGACTCC 420  
Qy 421 TCCCCGAGAGTGCACGACAGCTGACCCCCCCCCCTCCAGACAGGCCCAAGAGGAG 480  
Db 421 TCCCCGAGAGTGCACGACAGCTGACCCCCCCCCCTCCAGACAGGCCCAAGAGGAG 480  
Qy 481 GAGCTGAGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGCCGAACTCCATGTCC 540  
Db 481 GAGCTGAGCTGGCTGTGCTCTGGGGCGCTGTGAGCTCCGGCCGAACTCCATGTCC 540  
Qy 541 AGCCTGCCCCACACAGCAGCAGCAGCTACAGCTGGAGCCCGCTGGTGCACACCCCTG 600



541	Db		AGCTTGCACACACAGCACAGCAGCAGCTACAGCTGGACCCGCTGGTCACACCCGCTG	600
601	Qy	GGACCCACAAGCGTTT	TGGGGGCTCCGCCCAACAATCACCCAGGGCATCGTCTCTCCAG	660
601	Db	GGACCCACAAGCGTTT	TGGGGGCTCCGCCCAACAATCACCCAGGGCATCGTCTCTCCAG	660
661	Qy	GACAGCAACATGATGAGCT	GAAAGGCTCTGTCTTCTCCAGCGGAGTAGCAAGCTGGC	720
661	Db	GACAGCAACATGATGAGCT	GAAAGGCTCTGTCTTCTCCAGCGGAGTAGCAAGCTGGC	720
721	Qy	CACCTCGAAACAAGGCAGACAAGGGCCCTCGTGTGTCGCTCCGCCATCTCCACGGACGAG	780	
721	Db	CACCTCGAAACAAGGCAGACAAGGGCCCTCGTGTGTCGCTCCGCCATCTCCACGGACGAG	780	
781	Qy	TGCAGCATCCAGGAGCTGGACGAAGACTGTTGGAGAGGAGGGCCCTCCAGAAGCTG	840	
781	Db	TGCAGCATCCAGGAGCTGGACGAAGACTGTTGGAGAGGAGGGCCCTCCAGAAGCTG	840	
841	Qy	CAGCGCAGCTTTCAGGAGAGGAGCTG	CCCTCAGCCTGGCCTACGAGGAGCGCCCGCGG	900
841	Db	CAGCGCAGCTTTCAGGAGAGGAGCTG	CCCTCAGCCTGGCCTACGAGGAGCGCCCGCGG	900
901	Qy	CGCTGCAGGACAGAGCTGGAGGGCCCGGAGCCCAAAGGCGGCAACAAGCTCAAGCAGGCC	960	
901	Db	CGCTGCAGGACAGAGCTGGAGGGCCCGGAGCCCAAAGGCGGCAACAAGCTCAAGCAGGCC	960	
961	Qy	TGCAGAAGAGCCAGCGCGCGACGAGTCTTGCACTGCACTGCAAGTTCAGCAG	1020	
961	Db	TGCAGAAGAGCCAGCGCGCGACGAGTCTTGCACTGCAAGTTCAGCAG	1020	
1021	Qy	GAGAAGCGGCAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGACCTGCTCGAG	1080	
1021	Db	GAGAAGCGGCAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGACCTGCTCGAG	1080	
1081	Qy	ACCAAGCTCAGGTCTCTACGAGAGGAGGAAGACAGACTTCGGCCCGCGCTGGAGAGACC	1140	
1081	Db	ACCAAGCTCAGGTCTCTACGAGAGGAGGAAGACAGACTTCGGCCCGCGCTGGAGAGACC	1140	
1141	Qy	CAGTGGGAGGTGTCAGAGTCAAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG	1200	
1141	Db	CAGTGGGAGGTGTCAGAGTCAAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAG	1200	
1201	Qy	TCCACGACGAGGTGAACGCGCAAGGCTAGCGAGATCTCGGTCTCAAGGCACAGCTCAAG	1260	
1201	Db	TCCACGACGAGGTGAACGCGCAAGGCTAGCGAGATCTCGGTCTCAAGGCACAGCTCAAG	1260	
1261	Qy	GACACCGGGGCAAGCTGGAGGGCCTGGAGCTGAGGACCCAGGACCTGGAGGGCCCTG	1320	
1261	Db	GACACCGGGGCAAGCTGGAGGGCCTGGAGCTGAGGACCCAGGACCTGGAGGGCCCTG	1320	
1321	Qy	CGCACCAAGGGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGCAGC	1377	
1321	Db	CGCACCAAGGGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGCAGC	1377	
1378	Qy	GAGGCGGAGCTGCTGGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTCGG	1437	
1381	Db	ATGTACCGGGGAACAGCGCCTTGAAGAGGCCCTTCGACGAGCTGGCACGTGGGGACAGC	1440	
1438	Qy	GCCACGGCCGCTTGGCCCGCGACATAGGGCCGCCACCTTCCCCGAGGACGCT	1491	
1441	Db	GCCGGGAGCCCTTGGAGGTGACCTTGAAGGGGCTGACATCCCTACGAGGAC	1494	

## RESULT 7

AAS30637

ID AAS30637 standard; DNA: 2411 BP.

XX

AC AAS30637;

XX

DT 21-NOV-2001 (first entry)

XX

DE DNA encoding novel lung cancer antigen, Seq ID No 89.

XX	Human; lung cancer; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; autoimmune disease; rheumatoid arthritis; neoplasm; hyperproliferative disorder; cardiovascular disorder; angiogenesis; nervous system disorder; Alzheimer's disease; infection; skin aging; ocular disorder; wound healing; organ transplantation; ds.
OS	Homo sapiens.
XX	
XX	W0200155300-A2.
PN	
XX	
PD	02-AUG-2001.
XX	
XX	17-JAN-2001; 2001WO-US01238.
XX	
XX	31-JAN-2000; 2000US-0179065.
PR	
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
XX	Rosen CA, Barash SC, Ruben SM;
PI	
XX	WPI; 2001-465565/50.
DR	
XX	
XX	Isolated nucleic acid molecule encoding a lung cancer antigen is used in preventing, treating or ameliorating a medical condition -
PT	
XX	
PS	Disclosure; SEQ ID NO 89; 475pp; English.
XX	
CC	The invention relates to novel isolated lung cancer antigen polynucleotides (I) and polypeptides (II). (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are also used in diagnosing a pathological condition or susceptibility to a pathological condition, in particular, lung cancer. The antibodies to (II) can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and chemotaxis. AAS30580-AAS30685 represent novel human lung cancer antigen coding sequences, PCR primers and related sequences of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at: ftp.wipo.int/pub/published_pct_sequences.
XX	
XX	Sequence 2411 BP; 558 A; 681 C; 733 G; 439 T; 0 other;
QY	Query Match 45.2%; Score 808.8; DB 22; Length 2411;
DB	Best Local Similarity 99.8%; Pred. No. 3e-125; Indels 0; Gaps 0;
XX	Matches 810; Conservative 0; Mismatches 2;
QY	345 GGCGCTCCGAGAGGGGTGCAGTGAGGCCACAGCCTTTCAAGCCCTGTGCTGCCACGGTCAGG 404
DB	1306 GGCGCTCCGAGAGGGGTGCAGTGAGGCCACAGCCTTTCAAGCCCTGTGCTGCCACGGTCAGG 1365
QY	405 AGCCATCTTGCACTCTCTCCCGAGAGTGCCAGCCACCAAGCTGCACCCGCCCTCCAGA 464
DB	1366 AGCCATCTTGCACTCTCTCCCGAGAGTGCCAGCCACCAAGCTGCACCCGCCCTCCAGA 1425
QY	465 CRAAGCCCAAGGACGAGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTCTGCAGACTCCGG 524
DB	1426 CAAGCCCAAGGACGAGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTCTGCAGACTCCGG 1485
QY	525 CCGGAAGCTCATGTCTCCAGCGCTGCCACACACAGCAGCCAGCAGCTACAGCTGGACCC 584

Db	1486	CCGNACTCCATGCTCCAGCTTCCACACACAGACACAGACAGCTACCACTGGACCC	1545
Qy	585	GCTGGTCACACCGTGGAGCCACAAAGCGTTTGGGGGCTCCGCCACAAATCACC	644
Db	1546	GCTGGTCACACCGTGGAGCCACAAAGCGTTTGGGGGCTCCGCCACAAATCACC	1605
Qy	645	GGGATCGTCTCCAGGACACACATGATGAGCTGAAGGCTGTCTCTCCGACGG	704
Db	1606	GGGATCGTCTCCAGGACACACATGATGAGCTGAAGGCTGTCTCTCCGACGG	1665
Qy	705	AGGTAGCAAGCTGGGCCACTCCAAAGGAGAGCAAGGCGCCCTCGTGTCCGCTCCC	764
Db	1666	AGGTAGCAAGCTGGGCCACTCCAAAGGAGAGCAAGGCGCCCTCGTGTCCGCTCCC	1725
Qy	765	CATCTCCAGCAGAGTCCAGATCCAGGAGCTGGAGAGAGAGCTGTGGAGAGGAGG	824
Db	1726	CATCTCCAGCAGAGTCCAGATCCAGGAGCTGGAGAGAGAGCTGTGGAGAGGAGG	1785
Qy	825	CGCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGGCCTA	884
Db	1786	CGCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGGCCTA	1845
Qy	885	CCAGGAGCGCGCGGCGCTCCAGGAGCAGCTGGAGGCGCGGAGCCCAAGCGGCA	944
Db	1846	CCAGGAGCGCGCGGCGCTCCAGGAGCAGCTGGAGGCGCGGAGCCCAAGCGGCA	1905
Qy	945	CAAGCTCAAGCAGGCTCCAGAGAGCAGCGCGGCGCAGGCTTCGTGACCTGCGAGT	1004
Db	1906	CAAGCTCAAGCAGGCTCCAGAGAGCAGCGCGGCGCAGGCTTCGTGACCTGCGAGT	1965
Qy	1005	ACTCAGCTTCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTTCGAGAGCTCATGAGGA	1064
Db	1966	ACTCAGCTTCAGCAGGAGAGCGGAGCTCCGCGAGGAGCTTCGAGAGCTCATGAGGA	2025
Qy	1065	CGAGGACCTGTGGAGACCAAGCTCAGGCTTACGAGGAGAGAGCAGCTTCGGGCC	1124
Db	2026	CGAGGACCTGTGGAGACCAAGCTCAGGCTTACGAGGAGAGAGCAGCTTCGGGCC	2085
Qy	1125	CGCGTGGAGAGACCCAGTGGAGGTGTGCC	1156
Db	2086	CGCGTGGAGAGACCCAGTGGAGGTGTGCC	2117

RESULT 8

AA528699	ID	AA528699 standard; DNA; 2411 BP.
XX	AC	AA528699;
XX	AC	AA528699;
DT	07-NOV-2001	(first entry)
XX	DE	Genomic sequence #539 encoding for novel human respiratory antigen.
XX	DE	Human; respiratory antigen; respiratory disorder; throat disorder;
KW	KW	lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW	KW	anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW	KW	respiratory active; ds.
XX	OS	Homo sapiens.
XX	OS	WO200155448-A1.
PN	PD	02-AUG-2001.
XX	PF	17-JAN-2001; 2001WO-US01333.
XX	PF	31-JAN-2000; 2000US-0179065.
PR	PR	04-FEB-2000; 2000US-0180628.
PR	PR	24-FEB-2000; 2000US-0184664.
PR	PR	02-MAR-2000; 2000US-0186350.
PR	PR	16-MAR-2000; 2000US-0199874.
PR	PR	17-MAR-2000; 2000US-0190076.

PR	18-APR-2000;	2000US-0198123.
PR	19-MAY-2000;	2000US-0205515.
PR	07-JUN-2000;	2000US-0209467.
PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	07-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216880.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-JUL-2000;	2000US-0220964.
PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226688.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	06-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233083.
PR	14-SEP-2000;	2000US-0233064.
PR	21-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	25-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	26-SEP-2000;	2000US-0234998.
PR	27-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249267.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-476224/51.  
XX  
PT Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the respiratory system including respiratory  
PT cancers and also for testing and detection e.g. diagnosis -  
XX  
PS Disclosure; SED ID No 1133; 546pp; English.  
XX  
XX The present invention relates to the isolation of novel human  
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic  
CC sequences encoding for these polypeptides. The sequences of the

CC invention are useful for preventing, treating and/or prognosing  
CC disorders related to the respiratory system including throat  
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),  
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,  
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of  
CC the respiratory tissues e.g. lung cancer. the polynucleotide sequences  
CC of the invention are useful in gene therapy and antisense therapy.  
CC AAS28161-AAS28764 represent genomic sequences encoding for novel  
CC human respiratory antigens.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2411 BP; 558 A; 681 C; 733 G; 439 T; 0 other;  
Query Match 45.2%; Score 808.8; DB 22; Length 2411;  
Best Local Similarity 99.8%; Pred. No. 3e-125;  
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 345 GGGCTCCGAGAAGGGTGCAGTGAGGCCACACAGCCCTTCAAGCCTGTGTGTCACAGGTGAGG 404  
DB 1306 GGGCTCCGAGAAGGGTGCAGTGAGGCCACACAGCCCTTCAAGCCTGTGTGTCACAGGTGAGG 1365  
QY 405 AGCCATCTCTCACTCTCTCCCGGAGAGTGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 464  
DB 1366 AGCCATCTCTCACTCTCTCCCGGAGAGTGCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1425  
QY 465 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTGTGTCTCTGGGGCGCTGTGTGTGTGTGTGTGTGT 524  
DB 1426 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCTGTGTCTCTGGGGCGCTGTGTGTGTGTGTGTGTGT 1485  
QY 525 CCGGAACCTCCATGTCCAGCCTGCCACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 584  
DB 1486 CCGGAACCTCCATGTCCAGCCTGCCACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1545  
QY 585 GCTGTGTACACCCCGTGGGACCCCAAGCCGTTTTTGGGGGCTTCGCGCCACACATCACCCA 644  
DB 1546 GCTGTGTACACCCCGTGGGACCCCAAGCCGTTTTTGGGGGCTTCGCGCCACACATCACCCA 1605  
QY 645 GGGCATCGTCTCCAGGACAGCAACATGATGAGCTTGAAGGCTCTGTCTTCTCCGACGG 704  
DB 1606 GGGCATCGTCTCCAGGACAGCAACATGATGAGCTTGAAGGCTCTGTCTTCTCCGACGG 1665  
QY 705 AGGTAGCAAGCTGGGCCACTCGAACAAGGACAGCAAGGCCCTCGTGTGTGTGTGTGTGTGTGTGTGT 764  
DB 1666 AGGTAGCAAGCTGGGCCACTCGAACAAGGACAGCAAGGCCCTCGTGTGTGTGTGTGTGTGTGTGTGT 1725  
QY 765 CATCTCCAGGACAGGTGACGATCCAGGAGCTGGAGCAGAGCTGTTGGAGAGGGAGGG 824  
DB 1726 CATCTCCAGGACAGGTGACGATCCAGGAGCTGGAGCAGAGCTGTTGGAGAGGGAGGG 1785  
QY 825 CGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCTCCAGCCTGGCCCTA 884  
DB 1786 CGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCTCTCCAGCCTGGCCCTA 1845  
QY 885 CGAGGAGCGGCGCGCGCTGCAGGAGCAGCTTGAGGCGCCCGGAGCCCAAGCGCGCAA 944  
DB 1846 CGAGGAGCGGCGCGCGCTGCAGGAGCAGCTTGAGGCGCCCGGAGCCCAAGCGCGCAA 1905  
QY 945 CAAGCTCAAGCAGCGCTCGCAGAGAGCCAGCGCGCGCAGCAGGTCTGTGCACTTCAGGT 1004  
DB 1906 CAAGCTCAAGCAGCGCTCGCAGAGAGCCAGCGCGCGCAGCAGGTCTGTGCACTTCAGGT 1965  
QY 1005 ACTCAGCTTCAGCAGGAGAGCGCAGCTCCGCGCAGGAGCTCCGAGAGCCCTCATGAAGA 1064  
DB 1966 ACTCAGCTTCAGCAGGAGAGCGCAGCTCCGCGCAGGAGCTCCGAGAGCCCTCATGAAGA 2025  
QY 1065 GCAGGACCTGTGTGAGACCAAGCTCAGTCTCTAGCAGAGGAGAGAGCAGCTTCGGGCC 1124  
DB 2026 GCAGGACCTGTGTGAGACCAAGCTCAGTCTCTAGCAGAGGAGAGAGCAGCTTCGGGCC 2085  
QY 1125 CGCCTGGAGGAGACCCAGGTGGGAGGTGTGCC 1156  
|||||

```
Db      2086 CGCGCTGAGGAGAGCCAGTGGAGGTGAGGC 2117
RESULT 9
AAA64507
ID      AAA64507 standard; DNA; 9048 BP.
XX
AC      AAA64507;
XX
DT      02-JAN-2001 (first entry)
XX
DE      Nucleotide sequence comprising the human FEZ1 gene.
XX
KW      Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;
KW      tumour proliferation; tubulin; microtubule; protein EPI-gamma;
KW      tubulin polymerisation disorder; mitosis initiation; cell proliferation;
KW      cell growth; cell shape; cell rigidity; cell motility; DNA replication;
KW      tumorigenesis; tumour survival; metastasis; ss.
XX
OS      Homo sapiens.
XX
PN      WO2000050565-A2.
XX
PD      31-AUG-2000.
XX
PF      25-FEB-2000; 2000WO-US04950.
XX
PR      25-FEB-1999; 99US-0121537.
XX
PA      (UUYE-) UNIV JEFFERSON THOMAS.
XX
PI      Croce CM, Ishii H;
XX
WPI; 2000-558396/51.
XX
New polynucleotide homologous with a portion of one strand of the human
FEZ1 gene, useful for alleviating abnormal cell proliferation such as
cancer -
XX
Claim 2; Fig 5A; 255pp; English.
XX
The present sequence comprises the human FEZ1 gene. FEZ1 is a tumour
suppressor gene, located at chromosome location 8p22. Decreased or no
expression of FEZ1 is detected in a variety of cancer cells. Expression
of FEZ1 inhibits tumour growth and proliferation. FEZ1 also interacts
with tubulin, with microtubules, and with protein EPI-gamma.
XX
Post-translational phosphorylation and dephosphorylation modulates the
effect of the FEZ1 protein. Inhibitors of FEZ1 gene expression are
useful for inducing cells to proliferate. Compounds which modulate FEZ1
association with tubulin are useful for alleviating tubulin hyper- or
hypo- polymerisation disorders, such as those associated with aberrant
initiation of mitosis, modulation of the initiation and rate of cell
proliferation and cell growth, modulation of cell shape, cell rigidity,
cell motility, rate and stage of cellular DNA replication, intracellular
distribution of organelles, metastatic potential of cell and cellular
transformation from a non-cancerous to cancerous phenotype. Compounds
which modulate FEZ1 binding and phosphorylation are also useful for
alleviating a disorder, such as tumorigenesis, tumour survival, growth
and metastasis.
XX
SQ      Sequence 9048 BP; 2011 A; 2605 C; 2490 G; 1939 T; 3 other;
Query Match      45.1%; Score 807.2; DB 21; Length 9048;
Best Local Similarity 99.6%; Pred. No. 5.8e-125;
Matches 809; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
345 GGGCTCCGAGAGGGTGCAGTGAGGCCACAGCCTTCAAGCTGTGCTGCCACGGTCAGG 404
1646 GGGCTCCGAGAGGGTGCAGTGAGGCCACAGCCTTCAAGCTGTGCTGCCACGGTCAGG 1705
XX
405 AGCCATCTGCACTCTCTCCCGGAGAGTGCCAGCCACAGTCGACCCCGCCCTCCAGA 464
1706 AGCCATCTGCACTCTCTCCCGGAGAGTGCCAGCCACAGTCGACCCCGCCCTCCAGA 1765
XX
```





29-SEP-2000; 2000US-0236369.  
29-SEP-2000; 2000US-0236370.  
02-OCT-2000; 2000US-0236802.  
02-OCT-2000; 2000US-0237037.  
02-OCT-2000; 2000US-0237038.  
02-OCT-2000; 2000US-0237039.  
02-OCT-2000; 2000US-0237040.  
13-OCT-2000; 2000US-0239935.  
13-OCT-2000; 2000US-0239937.  
20-OCT-2000; 2000US-0240960.  
20-OCT-2000; 2000US-0241221.  
20-OCT-2000; 2000US-0241785.  
20-OCT-2000; 2000US-0241786.  
20-OCT-2000; 2000US-0241787.  
20-OCT-2000; 2000US-0241808.  
20-OCT-2000; 2000US-0241809.  
01-NOV-2000; 2000US-0241826.  
08-NOV-2000; 2000US-0244617.  
08-NOV-2000; 2000US-0246474.  
08-NOV-2000; 2000US-0246475.  
08-NOV-2000; 2000US-0246476.  
08-NOV-2000; 2000US-0246477.  
08-NOV-2000; 2000US-0246478.  
08-NOV-2000; 2000US-0246523.  
08-NOV-2000; 2000US-0246524.  
08-NOV-2000; 2000US-0246525.  
08-NOV-2000; 2000US-0246526.  
08-NOV-2000; 2000US-0246527.  
08-NOV-2000; 2000US-0246528.  
08-NOV-2000; 2000US-0246532.  
08-NOV-2000; 2000US-0246609.  
08-NOV-2000; 2000US-0246610.  
08-NOV-2000; 2000US-0246611.  
08-NOV-2000; 2000US-0246613.  
17-NOV-2000; 2000US-0249207.  
17-NOV-2000; 2000US-0249208.  
17-NOV-2000; 2000US-0249209.  
17-NOV-2000; 2000US-0249210.  
17-NOV-2000; 2000US-0249211.  
17-NOV-2000; 2000US-0249212.  
17-NOV-2000; 2000US-0249213.  
17-NOV-2000; 2000US-0249214.  
17-NOV-2000; 2000US-0249215.  
17-NOV-2000; 2000US-0249216.  
17-NOV-2000; 2000US-0249217.  
17-NOV-2000; 2000US-0249218.  
17-NOV-2000; 2000US-0249244.  
17-NOV-2000; 2000US-0249245.  
17-NOV-2000; 2000US-0249264.  
17-NOV-2000; 2000US-0249265.  
17-NOV-2000; 2000US-0249267.  
17-NOV-2000; 2000US-0249299.  
17-NOV-2000; 2000US-0249300.  
01-DEC-2000; 2000US-0250160.  
01-DEC-2000; 2000US-0250391.  
05-DEC-2000; 2000US-0251030.  
05-DEC-2000; 2000US-0251988.  
05-DEC-2000; 2000US-0256719.  
06-DEC-2000; 2000US-0251479.  
08-DEC-2000; 2000US-0251856.  
08-DEC-2000; 2000US-0251868.  
08-DEC-2000; 2000US-0251869.  
08-DEC-2000; 2000US-0251989.  
08-DEC-2000; 2000US-0251990.  
11-DEC-2000; 2000US-0254097.  
05-JAN-2001; 2001US-0259678.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-476224/51.  
P-PSDB; AAU17943.

XX Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the respiratory system including respiratory  
PT cancers and also for testing and detection e.g. diagnosis -  
XX  
PS Claim 4; SED ID No 269; 546pp; English.  
XX  
CC The present invention relates to the isolation of novel human  
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic  
CC sequences encoding for these polypeptides. The sequences of the  
CC invention are useful for preventing, treating and/or prognosing  
CC disorders related to the respiratory system including throat  
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),  
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,  
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of  
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences  
CC of the invention are useful in gene therapy and antisense therapy.  
CC AAS27869-AAS28159 encode for novel human respiratory antigens.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
SQ Sequence 560 BP; 120 A; 136 C; 161 G; 83 T; 0 other;  
Query Match 31.3%; Score 560; DB 22; Length 560;  
Best Local Similarity 100.0%; Pred. No. 4.3e-84;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGTAATGGGCTCCGAGAAGGTCAGTGAGGC 370  
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGTAATGGGCTCCGAGAAGGTCAGTGAGGC 60  
Qy 371 CCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGGAGCCATCTGCACCTCTCTCCCGGAGA 430  
Db . 61 CCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGGAGCCATCTGCACCTCTCTCCCGGAGA 120  
Qy 431 GTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGGAGCAGAGCTGAAGC 490  
Db 121 GTGCCAGCCACAGCTGCACCCCGCCCTCCAGACAAGCCCAAGGAGCAGAGCTGAAGC 180  
Qy 491 CTGGCCCTGTGCTCTGGGGGGCTGTGCAGACTCCGGCCGGAACTCCATGTCCAGCTGCCCA 550  
Db 181 CTGGCCCTGTGCTCTGGGGGGCTGTGCAGACTCCGGCCGGAACTCCATGTCCAGCTGCCCA 240  
Qy 551 CACACAGCACCAGCAGCAGCTACACAGCTGGACCCGCTGGTCCACACCCGTGGGACCCACAA 610  
Db 241 CACACAGCACCAGCAGCAGCTACACAGCTGGACCCGCTGGTCCACACCCGTGGGACCCACAA 300  
Qy 611 GCCGTTTTGGGGGCTCCGCCCAACAATCATCCAGGGCATCGTCTCCAGACAGCAACA 670  
Db 301 GCCGTTTTGGGGGCTCCGCCCAACAATCATCCAGGGCATCGTCTCCAGACAGCAACA 360  
Qy 671 TGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCCACTCGAACA 730  
Db 361 TGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCCACTCGAACA 420  
Qy 731 AGGCAGACAAGGGCCCTCGTGTCTCCGCTCCCATCTCCACGAGCAGGTGCAGCATCC 790  
Db 421 AGGCAGACAAGGGCCCTCGTGTCTCCGCTCCCATCTCCACGAGCAGGTGCAGCATCC 480  
Qy 791 AGGAGCTGGAGCAGAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGGGCAGCT 850  
Db 481 AGGAGCTGGAGCAGAAGCTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGGGCAGCT 540  
Qy 851 TTGAGGAGAGGAGCTTGCC 870  
Db 541 TTGAGGAGAGGAGCTTGCC 560  
RESULT 13  
ABQ54361  
ID ABQ54361 standard; cDNA; 1623 BP.  
XX

ABQ54361;  
 22-AUG-2002 (first entry)  
 Human ovarian antigen HELGN26 cDNA, SEQ ID NO:241.  
 Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 inflammatory condition; immune disorder; blood disorder;  
 cardiovascular disorder; respiratory disorder; neurological disorder;  
 gastrointestinal disorder; urinary system disorder; drug screening;  
 gene therapy; chromosome mapping; forensic analysis;  
 antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 antiinflammatory; gynaecological; reproductive; gene; ss.  
 Homo sapiens.  
 WO200200677-A1.  
 03-JAN-2002.  
 07-JUN-2001; 2001WO-US18569.  
 07-JUN-2000; 2000US-209467P.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Birse CE, Rosen CA;  
 WPI: 2002-147878/19.  
 P-PSDB; ABP41284.  
 Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 useful in the prevention, treatment and diagnosis of cancer (e.g.  
 ovarian cancer), immune disorders, cardiovascular disorders and  
 neurological diseases -  
 Claim 1; SEQ ID No 241; 2922pp; English.  
 The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 encompasses polypeptides 90% identical and polynucleotides 95% identical  
 to the sequences of the invention. The invention additionally relates to  
 recombinant vectors and host cells comprising human ovarian antigen  
 polynucleotides, antibodies against human ovarian antigens, and the use  
 of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 treating, prognosing or preventing various ovary and/or breast-related  
 disorders. Such conditions include ovarian cancer and breast cancer, and  
 metastatic tumours of ovarian or breast origin, reproductive system  
 disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 vaginitis), immune disorders (e.g., congenital and acquired  
 immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 respiratory disorders, neurological disorders, gastrointestinal disorders  
 and urinary system disorders. Ovarian antigen polypeptides and  
 polynucleotides may also be used in screening for compounds which  
 modulate ovarian antigen expression or activity. The polynucleotides may  
 further be used for gene therapy, chromosome mapping, in the  
 identification of individuals and in forensic analysis, and the  
 polypeptides may be used as food additives or to prepare antibodies  
 useful in disease diagnosis, drug targeting and phenotyping. The present  
 sequence represents cDNA encoding a human ovarian antigen of the  
 invention.  
 Note: The sequence data for this patent did not form part of the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences.  
 Sequence 1623 BP; 358 A; 457 C; 534 G; 268 T; 6 other;  
 Query Match 15.1%; Score 269.8; DB 24; Length 1623;  
 Best Local Similarity 60.2%; Pred No. 5e-36;  
 Matches 524; Conservative 0; Mismatches 304; Indels 43; Gaps 3;  
 Qy 962 CGCAGAGAGCCAGCGCGCAGCAGGTCTCTGACCTCTGACAGTCTGACAGTCTGACAGG 1021  
 Db 33 CACAGCGGGCACAGCGCGGCCAACAGNTGTGTGACGTGTGACAGTGTTCAGCTGACAGG 92  
 Qy 1022 AGAAGCGGAGCTCCGGCAGGAGCTCGAGAGCTCTATGAAGGAGCAGGACCTGCTGGAGA 1081  
 Db 93 AGAAGCGGCAATTGCAAGGACGACTTTCACAGCTGTCTGCAAGGAGCGCAACAGCTGGAGC 152  
 Qy 1082 CCAAGCTCAGGTCTCTAGAGAGGAGAGAACACAG-CTTCGGCCCGCGCTGGAGAGAC 1140  
 Db 153 GCGCTCGCCACCTTGGAGCGGGAGCAGCNGGGAGCTCGGGCGAGCTTGAAGAGACC 212  
 Qy 1141 CAGTGGAGGTGTGCCAGAAAGTCAAGCAGATCTCCCTCTGAGAGCAGCTGAAGAG 1200  
 Db 213 AAGTGGAGGTGTGCCAGAAATCAGGCGAGATCTCCCTGTGAAGCAGCAGCTGAAGAG 272  
 Qy 1201 TCCAGAGCGAGGTGAACGCCAAGGCTAGCAGATCTCTGGTCTCAAGGACACAGCTGAAG 1260  
 Db 273 TCTCAGCAGAGCTGTGTGCAGAAAGGCGAGGCTGTGTGGCTCTGCGGGTGGCGCTCGG 332  
 Qy 1261 GACACGCGGGCAGCTGGAGGGCTGTGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTG 1320  
 Db 333 GAGGCCGTGTACCTGCGGGTCACTGAGGGCGTCTGCGGGTCTACAGGAGCGGCC 392  
 Qy 1321 CGCACCAAGGCGCTGGAGCTGGAGGTCTGTGAGAAATGAGCTGACGCGCAAGAAAGAG 1380  
 Db 393 CGAGCTCGGGAGCTGGAGCTGGAGGAGCTGTTCAGGAGCTGACGAGCACACCGCCAGAA 452  
 Qy 1381 GCGAGCTGTCTGCGGAGAGAGTGAACCTGCTGAGCAGGAGCTGACGAGGCTGCGGG-- 1438  
 Db 453 GCTGAGCAGCTGCGGGAGAAAGCTGGGAGCTTGGATGCTGAGGCGCGCGGAGCTCGGGAG 512  
 Qy 1439 -----CCACGCGCGCCCTGGCGCGGACATG 1464  
 Db 513 CCCCTGTGCACCTGCCACCGCTGACCCATCTCTCTGCGAGAGATGATGAGGCCAAA 572  
 Qy 1465 GGGCGGCCACCTTCCCGGAGGAGCTCCCTGCCCTGAGCGGGAGCTGGAGCGGCTGCGG 1524  
 Db 573 GTGCAGCGGGCAGCAGCGCGGGTTGGGGGAGCTTGGCGGGCCAGGTGGAGCGGATTCGGG 632  
 Qy 1525 GCCAGCTGCGGGAGGAGCGGCAAGGCATACAGATGCTCTCGGCTTCCAGCATGAG 1584  
 Db 633 GTGAGCTGACGCGGGAGCGCGCGCGGKGTGAGGAGCAGCGGGGACAGCTTTGAGGGGAG 692  
 Qy 1585 CGGCTCTGTGTGGAAGGAGGAGAGAGTATTGATTCAGTACCAAGAAACAGCTGCAGCAG 1644  
 Db 693 CGGCTGCGCTGGCAGGAGAGAGAGAGAGAGAGTATCGCTACCAAGAGCAGCTGCAGCAG 752  
 Qy 1645 AGCTACCTGGCCATGTATACAGCGGAACAGCGCTGGAGAGAGCGCTTGCAGCAGCTGGCA 1704  
 Db 753 AACTATATCCAGATGTATACCGCGCAACCGGCGAGCTAGAGCAGGAGCTGCAGCAGCTCAGC 812  
 Qy 1705 CGTGGGAGACGCGCGGGAGCCCTTGGAGTTCACTGG-----AAGGGGCTGACATC 1758  
 Db 813 CTGGAGCTGGAGGCGCGGAGCTCGCTGACCTGGGCTGGCGGAGCAGGCGCCCTGTCATC 872  
 Qy 1759 CCCTACGAGGACATCATAGCCACTGAGATCT 1789  
 Db 873 TGCTGGAGGAGATCATCTGCTACTGAGATCT 903  
 RESULT 14  
 AAF15625  
 ID AAF15625 standard; cDNA; 1617 BP.  
 XX AAF15625;  
 AC AAF15625;  
 XX  
 DT 13-MAR-2001 (first entry)



XX DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:60.

XX DE Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;

KW vulnary; gastrointestinal; nephrotropic; antinefactive; gynaecological;

KW antibacterial; gene therapy; neural; immune; reproductive; renal;

KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease; ss.

OS Homo sapiens.

XX WO200055174-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05988.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (ROSE/) ROSEN C A.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2000-587513/55.

XX DR P-PSDB; AAB56422.

XX PT Prostate cancer associated gene sequences, referred to as prostate

PT cancer antigens, useful for treatment, prevention, and diagnosis of

PT disorders such as prostate cancer -

XX Claim 1; Page 657; 2338pp; English.

XX CC AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytostatic,

CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,

CC nephrotropic, antinefactive, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present

CC invention.

XX SQ Sequence 1617 BP; 355 A; 457 C; 532 G; 268 T; 5 other;

Query Match 14.4%; Score 258.2; DB 21; Length 1617;

Best Local Similarity 60.2%; Pred No. 4.1e-34;

Matches 523; Conservative 0; Mismatches 303; Indels 43; Gaps 4;

QY 962 CCAGAGAGCCAGCGCGCAGCAGGCTTCACCTGCAGTACTGCAGCTTCAGCAGG 1021

Db 33 CACAGCGGCGACAGCGGCGCCAAAGTGTGTCAGCTGCAGTGTCCAGCTGCA-CAGG 91

QY 1022 AGAAGCGGAGCTCCGGCAGGAGCTCGAGAGCTTCATGAAGCAGCAGCTTCGGAGA 1081

Db 92 AGAAGCGGCAATTGCGAGGACGACTTCGCACAGCTGTCTGAGGCGCGCAACAGCTGGAGC 151

QY 1082 CCAAGCTCAGGTCTTACAGAGGAGAGACCCAGCTTCGGCCCGCTGGAGGAGCCC 1141

Db 152 GCGCGTCCGCCACCTTGGAGCGGAGACAGCGGAGC-TCGGCGGAGGCTTCAGGAGCCA 210

QY 1142 AGTGGAGGTGTGCGAGAGTTCAGCGGAGATCTCCCTCTGAAGCAGCAGCTGAAGAGT 1201

Db 211 AGTGGAGGTGTGCGAGAGTTCAGCGGAGATCTCCCTCTGAAGCAGCAGCTGAAGAGT 270

QY 1202 CCCAGACGGAGGTGAACCCAGGCTAGCAGATCTTCGGGTCTCAAGGCACAGCTGAAGG 1261

Db 271 CTCAGGCAGAGCTGGTGCAGAAAGGCGCAGCTGGTGGCTCTCGCGGTGCGGTGCGG 330

QY 1262 ACAGCGCGGCAAGCTGGAGGCGCTGGAGCTGAGAGCCAGAGCTCTGAGGCGGCGCTTCG 1321

Db 331 AGGCGCGTGTACCTGCGGTGAGTGGGCGCTGCGGCGGCTTACAGAGGCGCGCC 390

QY 1322 GCACAAAGGCGCTGGAGCTGGAGTCTGTGAGATGAGTTCAGCGCAAGAAACGAGG 1381

Db 391 GAGCTCGGAGCTGGAGCTGGAAGCTGTTCCAGGAGCTGCAGCGACACCGCCAGGAAG 450

QY 1382 CGGAGCTGCTCGCGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGAGCTGCGGG--- 1438

Db 451 CTGAGCAGCTCGCGGAGAAAGCTGGGAGTTGGAATGCTGGCGCGCGGACTCCCGGAGC 510

QY 1439 -----CCAGGCGCGCTGCGCCCGCGACATGG 1465

Db 511 CCCCTGTGCCACCTGCCACCGCTGACCCATTCCTCTGCGAGAGAGTATGAGGCCAAG 570

QY 1466 GGCGCGCCACCTTCCCGAGAGACGTCCCTGTCAGCGGAGGAGTGGAGCGGTGCGGG 1525

Db 571 TGCAGCGGCGCAGCAGCGCGGCTTGGGGCGAGCTTGGGGCGCAGGTGGAGCGATTCGCGG 630

QY 1526 CCGAGCTCGCGGAGGCGGCAAGCCATGACCATGTCCTCGGGCTTCAGCATGAGC 1585

Db 631 TGGAGCTGCAGCGGAGCGCGCGGCTGAGGAGCAGCGGAGCAGCTTTGAGGGGAGC 690

QY 1586 GGCTCGTGTGGAAGGAGGAGAGAGAGAGGTGATTTCAGTACAGAAACAGCTGCAGCAGA 1645

Db 691 GGCTGGCTGCGAGCAGAGAGAGAGAGAGGTGATTTCAGTACAGAGAGAGCTGCAGCACA 750

QY 1646 GCTAGTGGCCATGTCACAGCGGAAACAGCGGCTTGGAGAGCGCCCTGCGAGCTGGCAC 1705

Db 751 ACTACATCCAGATGTACCGCGGCAACCGGCGAGCTAGAGCAGGAGCTGCAGCAGCTCAGCC 810

QY 1706 GTGGGAGCAGCGCGCGGAGCGCTTGGAGGTGACCTG-----GAAGGGGCTGACATCCC 1760

Db 811 TGGAGCTGGAGCGCGCGGAGCTTCGCTGACCTGGGCGCTTGGCGGAGCAGCGCCCTGCTG 870

QY 1761 CTACGAGCAGATCATAGCCACTGAGATCT 1789

Db 871 CCTGAGGAGATCACTGCTACTGAGATCT 899

RESULT 15

AAA64510

ID AAA64510 standard; cDNA; 404 BP.

XX AAA64510;

XX DT 02-JAN-2001 (first entry)

XX DE Nucleotide sequence of truncated FEZ1 transcript E16T8.

XX KW Human; FEZ1 gene; tumour suppressor gene; 8p22; cancer; tumour growth;

KW tumour proliferation; tubulin; microtubule; protein B1-gamma;

KW tubulin polymerisation disorder; mitosis initiation; cell proliferation;

KW cell growth; cell shape; cell rigidity; cell motility; DNA replication;

KW tumorigenesis; tumour survival; metastasis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1..231

FT /\*tag= a

FT /product= "truncated FEZ1"

XX PN WO200050565-A2.

XX PD 31-AUG-2000.

XX PF 25-FEB-2000; 2000WO-US04950.

XX PR 25-FEB-1999; 99US-0121537.

XX (UYJE-) UNIV JEFFERSON THOMAS.  
XX PA Croce CM, Iehii H;  
XX PI WPI; 2000-558396/51.  
XX DR P-PSDB; AAB08717.  
XX DR  
XX  
XX New polynucleotide homologous with a portion of one strand of the human  
XX FE21 gene, useful for alleviating abnormal cell proliferation such as  
XX cancer -  
XX  
XX Disclosure; Fig 5C; 255pp; English.  
XX  
XX The present sequence encodes a truncated human FE21 polypeptide. The  
XX encoding mRNA is transcribed by tumour cells. FE21 is a tumour  
XX suppressor gene, located at chromosome location 8p22. Decreased  
XX or no expression of FE21 is detected in a variety of cancer cells.  
XX Expression of FE21 inhibits tumour growth and proliferation. FE21  
XX also interacts with tubulin, with microtubules, and with protein  
XX EF1-gamma. Post-translational phosphorylation and dephosphorylation  
XX modulates the effect of the FE21 protein. Inhibitors of FE21 gene  
XX expression are useful for inducing cells to proliferate. Compounds  
XX which modulate FE21 association with tubulin are useful for alleviating  
XX tubulin hyper- or hypo- polymerisation disorders, such as those  
XX associated with aberrant initiation of mitosis, modulation of the  
XX initiation and rate of cell proliferation and cell growth, modulation of  
XX cell shape, cell rigidity, cell motility, rate and stage of cellular  
XX DNA replication, intracellular distribution of organelles, metastatic  
XX potential of cell and cellular transformation from a non-cancerous to  
XX cancerous phenotype. Compounds which modulate FE21 binding and  
XX phosphorylation are also useful for alleviating a disorder, such as  
XX tumorigenesis, tumour survival, growth and metastasis.  
XX  
XX Sequence 404 BP; 92 A; 119 C; 127 G; 66 T; 0 other;  
SQ

Query Match 13.9%; Score 249; DB 21; Length 404;  
Best Local Similarity 100.0%; Pred. No. 1.3e-32;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1543 CGGCAAGGCCATGACCCAGATGCTCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAG 1602  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
156 CGGCAAGGCCATGACCCAGATGCTCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAG 215  
QY 1603 GAGAGGAGGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAGAGCTAGTGGCCATGTAC 1662  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
216 GAGAGGAGGAGGTGATTTCAGTACCAGAAACAGCTGCAGCAGAGCTAGTGGCCATGTAC 275  
QY 1663 CAGCGGAACCCAGCGCTCGAGAGGCCCTTGACAGCTGGGACAGCGCCCGG 1722  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
276 CAGCGGAACCCAGCGCTTGAGAGGCCCTTGACAGCTGGGACAGCGCCCGG 335  
QY 1723 GAGCCCTTTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACT 1782  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
336 GAGCCCTTTGAGGTTGACCTGGAAGGGCTGACATCCCTACGAGGACATCATAGCCACT 395  
QY 1783 GAGATCTGA 1791  
Db |||||||||||  
396 GAGATCTGA 404

Search completed: June 14, 2003, 23:43:13  
Job time : 419.345 secs

Result No.	Score	Query		DB	ID	Description	
		Match	Length				
1	146	8.2	1926	4	US-09-249-585A-2	Sequence 2, Appli	
2	146	8.2	2580	3	US-09-050-861-2	Sequence 2, Appli	
3	146	8.2	2580	4	US-09-359-083-2	Sequence 2, Appli	
C	4	146	8.2	5452	2	US-09-130-114-1	Sequence 1, Appli
	5	146	8.2	9600	4	US-08-910-647-1	Sequence 1, Appli
	6	146	8.2	9600	4	US-09-620-925-1	Sequence 1, Appli
7	146	8.2	10596	1	US-07-884-811-15	Sequence 15, Appl	
8	146	8.2	10596	1	US-07-885-971-15	Sequence 15, Appl	
9	146	8.2	10596	1	US-08-087-783A-15	Sequence 15, Appl	
10	146	8.2	10596	1	US-08-194-080B-15	Sequence 15, Appl	
11	146	8.2	10596	2	US-08-194-087-15	Sequence 15, Appl	
12	146	8.2	10596	5	PCR-US93-04648-15	Sequence 15, Appl	
13	138.4	7.7	3489	2	US-08-728-323A-1	Sequence 1, Appli	
14	138.4	7.7	3489	4	US-09-298-568-1	Sequence 1, Appli	
C	15	138.4	7.7	32207	2	US-08-770-379-20	Sequence 20, Appl
	16	138.4	7.7	32207	4	US-08-777-379A-20	Sequence 20, Appl
	17	138.4	7.7	32207	4	US-09-230-371A-20	Sequence 20, Appl
C	18	125.2	7.0	9551	1	US-08-056-200-93	Sequence 93, Appl
	19	125.2	7.0	9551	2	US-08-800-644-93	Sequence 93, Appl
	20	113.8	6.4	2338	1	US-08-425-069-1	Sequence 1, Appli
21	113.8	6.4	2338	2	US-08-317-844B-1	Sequence 1, Appli	
22	112.6	6.3	2793	3	US-08-209-747-1	Sequence 1, Appli	
23	112.6	6.3	2793	1	US-08-458-298-1	Sequence 1, Appli	
24	103.2	5.8	1995	1	US-08-425-069-3	Sequence 3, Appli	
25	103.2	5.8	1995	2	US-08-317-844B-3	Sequence 3, Appli	
26	92.2	5.1	2830	2	US-09-010-928B-1	Sequence 1, Appli	
27	89	5.0	1926	4	US-09-249-585A-4	Sequence 4, Appli	

QY 1367 GCAGAGAACGAGCGGAGCTGCTCGGAGAGAGGTGAACCTGCTGAGCAGAGCTGC 1426  
DB 621 AGGAGGGGCGAGAGGGGCGAGGAGCAGGAGGGGCGAGGAGGAGGGGCGAGGAGG 680  
QY 1427 AGGAGCTCGGGGCGGAGCGGCGCTTGGCCCGGCGACATGGGGCCGCCACCTTCCCGAGG 1486  
DB 681 AGGAGGGGCGAGGAGCAGGAGGGGCGAGGAGCAGGAGGGGCGAGGAGGGGCGAGGAG 740  
QY 1487 AGCTCCCTGCTGAGCGGAGCTGAGCGGCTGCGGGCCGAGCTGCGGGGCGAGCGGC 1546  
DB 741 AGGAGGGGCGAGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 800  
QY 1547 AAGGCCATGACAGATGCTTCCGGCTTCCAGCATGAGCGCTGCTGTGGAAGAGGAGA 1606  
DB 801 AGGAGCGAGGAGCAGGAGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAG 860  
QY 1607 AGGAGAGGCTGATTCAGTACCAGAACAGCTGCGAGCAGAGCTAGCTGCCCATGTACCAGC 1666  
DB 861 AGGAGCAGGAGGGGCGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 920  
QY 1667 GGAACCGGCTGAGAGAGGGCCCTGCGAGAGCTGCGAGCTGCGGAGCAGCGCCGGGAGC 1726  
DB 921 GCGAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 980  
QY 1727 CCTTGGAGCTTGACCTGGAAGGGGCTGACATCCCTACGAGG 1768  
DB 981 AGGCCGGGCTGAGGAGGAGCTGAGGAGGCGGGGCTGAGGAGG 1022

## RESULT 2

US-09-050-863-2  
; Sequence 2, Application US/09050863  
; Patent No. 611411  
; GENERAL INFORMATION:  
; APPLICANT: Lao, Ying  
; APPLICANT: Hiang, Betty  
; APPLICANT: Payan, Don  
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/050,863  
; FILING DATE: 30-MAR-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silva, Robin M.  
; REGISTRATION NUMBER: 38,304  
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 949-8711  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2580 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA  
US-09-050-863-2

Query Match 8.2%; Score 146; DB 3; Length 2580;

Best Local Similarity 49.5%; Pred. No. 2.5e-19;  
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;  
QY 1007 TGAGAGTTACGAGGAGAACCGGAGCTCCGGGAGAGCTCGAGAGCTCATGAGAGC 1066  
DB 644 TGAACAGGAGCAGGAGCAGGAGCGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGG 703  
QY 1067 AGGACCTGCTGGAGACCAAGCTCAGTCTACGAGAGGAGAGACACAGCTTCGGCCCCG 1126  
DB 704 AGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 763  
QY 1127 CGTTGAGGAGAGACCCAGTGGGAGGTGTCCAGAGTCCAGCGAGATCTCCCTCTCCTAAGC 1186  
DB 764 AGGAGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 823  
QY 1187 AGCAGCTGAAGAGATGCCAGCAGGAGGTGAACCAAGGCTAGCGAGATCTCTGGGTCTCA 1246  
DB 824 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 883  
QY 1247 AGCAGAGCTGAAGGACACCGCGGGGCAAGCTGGAGGGGCTGGAGCTGAGGAGCCAGGACC 1306  
DB 884 AGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 943  
QY 1307 TGAGAGGCGGCTCGCGCACCAAGGGCTTGAGGCTTGAGGCTCTGTGAGAAATGAGCTGAGC 1366  
DB 944 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1003  
QY 1367 GCAAGAAAGCAGAGGCGGAGCTGCTGCGGAGAGAGGTGAACCTGCTGAGCAGGAGCTGC 1426  
DB 1004 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1063  
QY 1427 AGGAGCTCGGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1486  
DB 1064 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1123  
QY 1487 ACGTCCCTGCTCGAGCGGAGCTGAGCGGCTGCGGGGCGGAGCTGCGGAGGAGGAGGAGG 1546  
DB 1124 AGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1183  
QY 1547 AAGCCATGACAGATGCTCTCGGGCTTCCAGCATGAGCGGCTGCTGTGAGAGGAGAGA 1606  
DB 1184 AGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1243  
QY 1607 AGGAGAGGCTGATTTCAGTACCAAGAACAGCTGCGAGCAGAGCTAGCTGGCCATGTACCAGC 1666  
DB 1244 AGGAGCAGGAGGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1303  
QY 1667 GGAACCGGCTTGAGAGAGGCGCTGCGAGCAGCTGCGACCTGGGAGCAGCGCCGGGAGC 1726  
DB 1304 GGCAGGAGGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1363  
QY 1727 CCTTGGAGTTGACCTGGAAGGGCTGACATCCCTACGAGG 1768  
DB 1364 AGGCCGGGCTGAGGAGGAGCAGTGGAGGCGGGGTCGAGGAGG 1405

## RESULT 3

US-09-359-081-2  
; Sequence 2, Application US/09359081  
; Patent No. 6316223  
; GENERAL INFORMATION:  
; APPLICANT: Lao, Ying  
; APPLICANT: Hiang, Betty  
; APPLICANT: Payan, Don  
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA

ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/359,081  
FILING DATE: 22-Jul-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/050,863  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-8711  
TELEFAX: (415) 949-8711  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2580 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-359-081-2

Query Match 8.2%; Score 146; DB 4; Length 2580;  
Best Local Similarity 49.5%; Pred. No. 2.5e-19;  
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;  
QY 1007 TGCAGCTTCAGCAGGAGCGGCGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGC 1066  
DB 644 TGGAAACAGGAGCAGGAGGAGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 703  
QY 1067 AGGACCTGCTGAGACCAAGCTCAGGTCTACGAGAGGAGAGACCAAGCTTCGGCCCGC 1126  
DB 704 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 763  
QY 1127 CGCTGGAGAGACCCAGTGGAGGTGTCCAGAACTCAGGCGAGATCTCCCTCTGAAGC 1186  
DB 764 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 823  
QY 1187 AGCAGCTGAAGAGTCCAGAGGAGGTGAACCCAGGCTAGCAGATCTTGGGTCTCA 1246  
DB 824 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 883  
QY 1247 AGGCACAGCTGAAGACACCGCGGGGCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACC 1306  
DB 884 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 943  
QY 1307 TGGAGGGCCCTCGCCACCAAGGCGCTTGAGCTGGAGTCTGTGAGAAATGAGTGCAGC 1366  
DB 944 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1003  
QY 1367 GCAAGAAGCAGGCGGAGTGTCTCGGGAGAGGTGAACCTGTGAGCAGGAGTGC 1426  
DB 1004 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1063  
QY 1427 AGGAGCTGGGGCCAGGCGCGCTTGGCCCGGACATGGGGCCGCCACCTTCCCGGAGG 1486  
DB 1064 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1123  
QY 1487 AGCTCCCTCGCTGAGCGGAGCTTGGAGCGGCTCGCGGCCAGCTCGGGGAGGAGCGGC 1546  
DB 1124 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1183  
QY 1547 AAGGCCATGACCATGTCTCTCGGCTTCACAGCATGAGGGCTGTGTGGAAGGAGGAGA 1606  
DB 1184 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1243

QY 1607 AGGAGAGGTGATTTCAGTACCAAGAAACAGCTGCAGCAGAGCTAGCTGTGCCATGTACCAGC 1666  
DB 1244 AGGAGCAGGAGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1303  
QY 1667 GGAACACAGCGCTGGAGAGGCGCTTCAGCAGCTGGACAGCTGGGAGCAGCCCGGGAGC 1726  
DB 1304 GGCAGGAGGCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1363  
QY 1727 CTTGGAGTTGACTGGAAGGGCTGACATCCCTACGAGG 1768  
DB 1364 AGGCGGGCTCAGGAGGAGCAGTGGAGCCGGGTGAGAGG 1405

## RESULT 4

US-09-130-114-1/c  
; Sequence 1, Application US/09130114  
; Patent No. 5978807  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert A.  
; APPLICANT: Damaj, Bassam B.  
; APPLICANT: Robbins, Alan K.  
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes  
; TITLE OF INVENTION: From Multiple Transfected Episomes  
; FILE REFERENCE: 0867/1D903US1  
; CURRENT APPLICATION NUMBER: US/09/130.114  
; CURRENT FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 5452  
; TYPE: DNA  
; ORGANISM: VEBNA  
US-09-130-114-1

Query Match 8.2%; Score 146; DB 2; Length 5452;  
Best Local Similarity 49.5%; Pred. No. 2.9e-19;  
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;  
QY 1007 TGCAGCTTCAGCAGGAGCGGCGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGC 1066  
DB 2161 TGGAAACAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2102  
QY 1067 AGGACCTGCTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGGCCCGC 1126  
DB 2101 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2042  
QY 1127 CGCTGGAGAGACCCAGTGGAGGTGTGCCAGAACTCAGGCGAGATCTCCCTCTGAAGC 1186  
DB 2041 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1982  
QY 1187 AGCAGCTGAAGAGTCCAGAGCGGAGGTGAACCCAGGCTAGCGAGATCTTGGGTCTCA 1246  
DB 1981 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1922  
QY 1247 AGGCACAGCTGAAGACACACCGCGGGGCAAGCTGGAGGGCTTGGAGCTGAGGACCCAGGACC 1306  
DB 1921 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1862  
QY 1307 TGGAGGGCGCCCTCGCCACCAAGGCGCTTGGAGCTTGGAGGTCTGTGAGAAATGAGTGCAGC 1366  
DB 1861 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1802  
QY 1367 GCAAGAAGCAGGCGGAGCTGTGCGGGAGAGGTGAACCTGTGAGCAGGAGTGC 1426  
DB 1801 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1742  
QY 1427 AGGAGCTGGGGCCAGGCGCGCTTGGCCCGGACATGGGGCCGCCACCTTCCCGGAGG 1486  
DB 1741 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1682  
QY 1487 ACCTCTCTGCTTCAGCGGAGGTGAGAGCGGCTTCGCGGCCCGAGCTGCGGAGGAGCGGC 1546



NAME: Fujita, Sharon M.  
REGISTRATION NUMBER: 38,459  
REFERENCE/DOCKET NUMBER: 1218.002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 923-2706  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9600 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-620-925-1

Query Match 8.2%; Score 146; DB 4; Length 9600;  
Best Local Similarity 49.5%; Pred. No. 3.2e-19;  
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

Qy 1007 TGCAGCTTCAGCAGGAGGAGGCGAGCTCCGGCAGGAGCTCCAGAGCTCATGAAGGAGC 1066  
Db 690 TGGACAGCAGCAGCAGCAGGCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 749  
Qy 1067 AGGACCTCTCGAGACCAAGCTCAGTCTCTAGAGAGGAGAGACCAAGCTTCGGCCCGC 1126  
Db 750 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 809  
Qy 1127 CGCTGGAGGAGACCCAGTGGAGGTGTCCAGAGTCAAGGAGGAGATCTCCCTCTGAAGC 1186  
Db 810 AGGAGGCGCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 869  
Qy 1187 AGCAGCTGAAGAGAGTCCAGCAGGAGGTGAACCCAGGCTAGCAGATCTCTGGGTCTCA 1246  
Db 870 GGCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 929  
Qy 1247 AGGCACAGCTGAAGACACGCGGGCAAGCTGGAGGGCTCGAGCTGAGGACCCAGGACC 1306  
Db 930 AGGGCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 989  
Qy 1307 TGGAGGGCGCTTCGGCACCAAGGCTTGGAGCTGGAGGTCTGTGAGATGAGCTGCAGC 1366  
Db 990 AGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1049  
Qy 1367 GCAAGAGAACAGCAGGAGCTGTCTCGGAGAGAGGTGAACCTGTCTGAGCAGGAGCTGC 1426  
Db 1050 AGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1109  
Qy 1427 AGGAGCTCGGGCCAGCGCGCTTGGCCCGCAGCATGGGGCCGCCACCTTCCCGGAGG 1486  
Db 1110 AGGAGGGCGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1169  
Qy 1487 ACCTCCCTGCTTCAGCGGGAGCTTGGAGCGGCTCGGGCCGAGCTCGGGAGGAGCGGC 1546  
Db 1170 AGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1229  
Qy 1547 AAGGCCATGACAGATGTCTCGGCTTCCAGCATGAGCGGTCTGTGTGAAGAGGAGAGA 1606  
Db 1230 AGGGCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1289  
Qy 1607 AGGAGAGGTGATTTCAGTACCAGAAACAGCTGACAGCAGAGTACGTGCGCCATGTACCAGC 1666  
Db 1290 AGGAGCAGCGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1349  
Qy 1667 GGACACCGCTTGAGAGGCGCTTCAGCAGAGCTGGACGTGGGACAGCGCCCGGGAGC 1726  
Db 1350 GGCAGGAGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1409  
Qy 1727 CCTTGGAGGTTGACCTGGAAGGGGCTGACATCCCTTACGAGG 1768  
Db 1410 AGGCCGGGTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1451

RESULT 7  
US-07-884-811-15  
; Sequence 15, Application US/07884811  
; Patent No. 5316921  
; GENERAL INFORMATION:  
; APLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.  
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/884,811  
; FILING DATE: 19920518  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: 755.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/952-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10596 bases  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-07-884-811-15

Query Match 8.2%; Score 146; DB 1; Length 10596;  
Best Local Similarity 49.5%; Pred. No. 3.2e-19;  
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;

Qy 1007 TGCAGCTTCAGCAGGAGGAGGCGAGCTCCGGCAGGAGCTCGAGAGCTCATGAAGGAGC 1066  
Db 2224 TGGAAACAGGAGCAGGAGCAGGAGCGGGAGGGGCGAGGAGGAGGAGGAGGAGGAGG 2283  
Qy 1067 AGGACCTGTGGAGACCAAGCTCAGTCTCTACGAGGAGGAGAACACAGCTTCGGCCCGC 1126  
Db 2284 AGGGCAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2343  
Qy 1127 CGCTGGAGGAGACCCAGTGGAGGTGTGCCAGAGTCAAGGAGATCTCCCTCTCTGAAGC 1186  
Db 2344 AGGAGGGCGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2403  
Qy 1187 AGCAGCTGAAGGAGTCCAGCAGCAGGAGGTGAACCGCAAGGCTAGCGAGATCTCGGTCTCA 1246  
Db 2404 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2463  
Qy 1247 AGGCACAGCTGAAGGACACCGCGGGCAAGCTTGGAGGGCTTGGAGCTGAGGACCCAGGACC 1306  
Db 2464 AGGGCAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2523  
Qy 1307 TGGAGGGCGCCCTCGGCACCAAGGGCTTGGAGCTTGGAGGTCTGTGAGAAATGAGCTGCAGC 1366  
Db 2524 AGGAGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2583  
Qy 1367 GCAAGAGAACAGGCGGAGCTGTGCGGGAGAGAGGTGAACCTCTGTGAGCAGGAGCTGC 1426

2584	AGGAGGGGCAGGAGGGGCGCAGGACAGAGAGGGGCGAGGACAGGAGGAGGGGCGAGGAGGGGC	2643
1427	AGGAGCTGCGGGGCCAGGCGCCCTGCGCCCGCGACATGCGGGCCGCCCCACCTTCCCCGAGG	1486
2644	AGAGGGGCAGGACGAGAGGGGCGAGGACAGAGGGGCGAGGACGAGGAGGGGCAGGAGC	2703
1487	ACGTCCCTGCCCCCTGCAGCGGAGCTGGAGCGGCTGCGGGCCGAGCTCGGGAGGAGCGGC	1546
2704	AGGAGGGGCAGGAGGGGCGAGGACGAGAGGGGCGAGGAGGGGCGAGGACGAGGAGGGCAGG	2763
1547	AAGGCCATGACAGATGCTCCGGCTTCACGACATGAGCGGCTCGTGTGAGAGGAGAGA	1606
2764	AGGGGCAGGACAGGAGGAGGGGCAGAGGGGCGAGGAGGGGCGAGGAGGAGGGGCGAGGAGGGC	2823
1607	AGGAGAGGTGATTTCAGTACCAAGAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAGC	1666
2824	AGGAGCAGGAGGGGCGAGGGGCGAGGACAGGAGGGGCGAGGAGGGGCGAGGACGAGGAGG	2883
1667	GGAAACAGCGCCCTGGAGAGGCCCTGCGAGCAGCTGGCACCGTGGGGGACAGCCCGGGGAGC	1726
2884	GGCAGGAGGGCAGGACAGGAGGAGGGGCGAGGACGAGGAGGGGCGAGGACGAGGAGGTGG	2943
1727	CCTTGGAGGTGACCTCGAAGGGGCTGACATCCCTACGAGG	1788
2944	AGGCCGGGGTCGAGGAGGGCAGTGAGGCGCGGGGTCGAGGAGG	2985

## RESULT 8

```

US-07-885-971-15
: Sequence 15, Application US/07985971
: Patent No. 5328837
: GENERAL INFORMATION:
: APPLICANT: GoDowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
: TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genetech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,971
FILING DATE: 19920518
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

```

US-07-885-971-15

```
Query Match      8.2%; Score 146; DB 1; Length 10596;
Best Local Similarity 49.5%; Pred. NO. 3.2e-19;
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;
```

1007	Qy	TCGACGCTTCAGCAGGAGAACCGCAGCTCCGCAGCAGAGCTCGAGACCTCATGAAGGAGC	1066
2224	Db	TGAAACAGGAGCAGGACAGGAGCGGAGGGGCAGGAGCAGGAGGGGCAGGACGACGAGG	2283
1067	Qy	AGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGAGAGACACAGCTTCGGGCCCG	1126
2284	Db	AGGGGCAGGACAGGAGGAGGGGCAGAGGGGCAGGAGGGGCAGGAGGGGCAGGACGAGG	2343
1127	Qy	CGCTGGAGGAGACCCAGTCGGGAGGCTGTGCAGAAAGTCAGGCGCAGATCTCCCTCCTGAAGC	1186
2344	Db	AGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGACGAGGAGG	2403
1187	Qy	AGCAGCTGAAGAGTCCCAGACGAGAGGTGAACGCCAAGGCTAGCGAGATCTCTGGTCTCA	1246
2404	Db	GGCAGGACGAGGAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGCAGG	2463
1247	Qy	AGGCACAGCTGAAGGACACGCGGGGCAGCTGGAGGGCCTGGAGCTTGAGACCCAGGACC	1306
2464	Db	AGGGGCAGGACAGGAGGAGGGGCAGGACGAGGAGGAGGGGCAGGAGGGGCAGGACGAGG	2523
1307	Qy	TGAGGGCGCCCTGCCCACCAAGGGCCTTGAGCTGAGAGTCTGTGAGAATGAGCTGCAGC	1366
2524	Db	AGGAGGGCAGGAGGGGCAGGAGGGGCAGGACAGGAGGAGGGGCAGGACGAGGAGGGGC	2583
1367	Qy	GCAAGAAAGAACAGGCGGAGTGTCTCGGGAGAGAGTGTAACCTGTCTGGAGCAGGAGCTGC	1426
2584	Db	AGGAGGGCAGGAGGGGCAGGACAGGAGGGGCAGGAGCAGGAGGAGGAGGGGCAGGAGGGGC	2643
1427	Qy	AGGAGCTGGGGCCAGGCGCCCTTGGCCCGGCACATGGGGCCGCCACCTTCCCGCAGG	1486
2644	Db	AGGAGGGCAGGACGAGGAGGGGCAGGACGAGGAGGGGCAGGAGGAGGAGGGGCAGGAGG	2703
1487	Qy	ACGTCCTCTGCTGCACGCGGAGCTCGAGCGGCTGCGGGCCAGCTGCGGGAGGAGCGGC	1546
2704	Db	AGGAGGGCAGGAGGGGCAGGACAGGAGGGGCAGGAGGGGCAGGAGGAGGAGGGGCAGG	2763
1547	Qy	AAGGCCATGACCCAGATGTCTCTGGGTTCAGCATGAGCGGCTGTGTGTGAAGGAGGAGA	1606
2764	Db	AGGGGCAGGACAGGAGGAGGGGCAGGAGGGGCAGGAGCAGGAGGAGGGGCAGGAGGGGC	2823
1607	Qy	AGGAGAGGTGATTCAGTACCAAGAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAGC	1666
2824	Db	AGGACAGGAGGGGCAGGAGGGGCAGGACGAGGAGGGGCAGGAGGGGCAGGACGAGGAGG	2883
1667	Qy	GGAAACAGCGCCTTGAGAGAAAGCCCTGCAGCAGCTGTGCACTGTGGGACAGCGCCGGGAGC	1726
2884	Db	GGCAGGAGGGCAGGACGAGGAGGGGCAGGAGGGGCAGGAGGGGCAGGAGGAGGAGG	2943
1727	Qy	CCTTGGAGTTGACCTGGAGGGGCTGTGACATCCCTTACGAGG	1768
2944	Db	AGGCCGGGTCGAGGAGGACAGTGGAGCGCGGGTCTCAGGAGG	2985

## RESULTS

RESOLUT 9  
US-08-087-783A-15  
; Sequence 15, Application US/08087783A  
; Patent No. 5547856  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.  
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS

ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS



Db	2824	AGAGCAGGAGGGCGCAGGAGGGCGAGGACGACGAGAGGGGCGAGGAGGGCGCAGGAGG	2888
Qy	1667	GGAAACAGGCGCCTGGAGAGAGGGCCCTTCAGCAGCTGGCACCTGGGGACAGCGCGCGGGGAGC	1726
Db	2884	GCAGGAGGGCGCAGGACGAGGAGGGCGCAGGAGGGCGCAGGAGGGCGCAGGAGGAGTGG	2943
Qy	1727	CCTTGGAGGTTGACCTGGAGGGGCTGACATCCCTTACGAGG	1768
Db	2944	AGGCCGGGGTCGAGGAGGAGCTGGAGGCGCGGGTCCAGGAGG	2985
RESULT 10			
US-08-194-088B-15			
; Sequence 15, Application US/08194088B			
; Patent No. 5580963			
; GENERAL INFORMATION:			
; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.			
; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Genentech, Inc.			
; STREET: 460 Point San Bruno Blvd			
; CITY: South San Francisco			
; STATE: California			
; COUNTRY: USA			
; ZIP: 94080			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: patin (Genentech)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/194,088B			
; FILING DATE: 09-FEB-1994			
; CLASSIFICATION: 530			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 07/884811			
; FILING DATE: 18-MAY-1992			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Gallegos, R. Thomas			
; REGISTRATION NUMBER: 32,692			
; REFERENCE/DOCKET NUMBER: 75SD1			
; TELEPHONE: 415/225-2614			
; TELEFAX: 415/952-9881			
; TELEX: 910/371-7168			
; INFORMATION FOR SEQ ID NO: 15:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 10596 bases			
; TYPE: nucleic acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
US-08-194-088B-15			
Query Match 8.2%; Score 146; DB 1; Length 10596;			
Best Local Similarity 49.5%; Pred. No. 3.2e-19;			
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;			
Qy	1007	TGCAGCTTCAGCAGGAGAGCGCGCAGCTCCGCGAGAGCTCAGAGCCTCATGAAGGAGC	1066
Db	2224	TGGAAACAGGACGAGGACGAGGCGGAGGGGCGAGGAGCGGAGGCGGCGAGGAGGAGG	2283
Qy	1067	AGGACTGCTGGAGACCAAGCTCAGGTCTCTACGAGGCGGAGAGACCAGCTTCGGGCCCG	1126
Db	2284	AGGGGCGAGGACGAGGAGGAGGGGCGAGGAGGGGCGAGGAGGGGCGAGGAGGAGCAGG	2343
Qy	1127	CGCTGGAGGAGACCCAGTCGGAGGTTGTGCGAGAAGTCCAGGCGAGATCTCCCTCCTGAAGC	1186
Db	2344	AGGAGGGCGAGGACGAGCAGGAGGAGGGGCGAGAGGGGCGAGGCGGCGAGGAGGAGG	2403
Qy	1187	AGCAGCTGAAGGAGTCCCGACAGCGGAGGTGAACCCCAAGGCTAGCGAGATCTCTGGGTCTCA	1246
Db	2404	GGCAGGACGAGGAGGAGGGGCGAGGAGGGGCGAGGAGCAGGAGGAGGGGCGAGGAGGGG	2463



ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/04648  
FILING DATE: 19930517  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/884811  
FILING DATE: 18-MAY-92  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/885971  
FILING DATE: 18-MAY-92  
ATTORNEY/AGENT INFORMATION:  
NAME: Dregler, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 755,779P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10596 bases  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US93-04648-15

Query Match 8.2%; Score 146; DB 5; Length 10596;  
Best Local Similarity 49.5%; Pred. No. 3.2e-19;  
Matches 377; Conservative 0; Mismatches 385; Indels 0; Gaps 0;  
Qy 1007 TGCAGCTTCAGCAGGAGGCGGAGCTCCGGCAGGAGCTCGAGAGCTCATGAGGAGC 1066  
Db 2224 TCGAACAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2283  
Qy 1067 AGGACCTCTCGAGACCAAGCTCAGCTCTACGAGAGGAGGAGGAGGAGGAGGAGG 1126  
Db 2284 AGGGCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2343  
Qy 1127 CGCTGGAGGAGACCCAGTGGGAGGTGTGCCAGAAAGTCAGGCGAGATCTCCCTCTGAAGC 1186  
Db 2344 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2403  
Qy 1187 AGCAGCTGAAGAGTCCAGACGAGGAGTGAAGCCAGGCTAGGAGATCTGGGTCTCA 1246  
Db 2404 GGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2463  
Qy 1247 AGGCACAGCTGAAGACACGCGGGGCAAGCTGGAGGGCTTGAGCTGAGGAGCCAGGAGC 1306  
Db 2464 AGGGCAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2523  
Qy 1307 TGGAGGGCCCTCGCGCACCAAGGCGCTTGGAGCTTGGAGGTCTGTGAGAAATGAGTTCGAGC 1366  
Db 2524 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2583  
Qy 1367 CCAAGAGACAGGAGGAGGAGTGTGCGGAGAGAGGTGAACCTGCTGAGCAGGAGCTGC 1426  
Db 2584 AGGAGGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2643  
Qy 1427 AGGAGCTCGGGGCCAGGCGCCCTTGGCCGACATGGGGCCGCCACCTTCCCGGAGG 1486  
Db 2644 AGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2703

Qy 1487 ACGTCCCTGCGCTGACAGCGGAGTGTGAGCGGTGCGGGCCGAGCTCGGGAGGAGCGGC 1546  
Db 2704 AGGAGGGGAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2763  
Qy 1547 AAGCCATGACCATGCTCTCGGCTTCAGCATGAGCGGCTCGTGTGGAGGAGGAGA 1606  
Db 2764 AGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2823  
Qy 1607 AGGAGAGGTGATTTCAGTACCCAGAAACAGCTGTCAGCAGAGAGTACGTGGCCATGTACCAAGC 1666  
Db 2824 AGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2883  
Qy 1667 GGAACAGCGCTTGGAGAAAGCCCTGCGAGCAGCTGGCACTGGGGACAGCCCGGGAGC 1726  
Db 2884 GGCAGGAGGGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2943  
Qy 1727 CTTGGAGGTGACCTGGAAGGGGCTGACATCCCTACGAGG 1768  
Db 2944 AGCCCGGGGTCGAGGAGGAGGAGTGGAGGCCCGGGGTCGAGGAGG 2985

RESULT 13  
US-08-728-323A-1  
; Sequence 1, Application US/08728323A  
; Patent No. 5948676  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA  
; TITLE OF INVENTION: Encoding Same And Uses Thereof  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/728,323A  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-278-0400  
; TELEFAX: 212-391-0525  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3489 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..3489  
US-08-728-323A-1

Query Match 7.7%; Score 138.4; DB 2; Length 3489;  
Best Local Similarity 47.0%; Pred. No. 7.3e-18;  
Matches 461; Conservative 0; Mismatches 516; Indels 3; Gaps 1;

Qy	790	CAGGAGCTGGAGCAGAAAGCTGTTGGAGAGGGAGGGCGCTCCAGAAAGTGTGCAGCGCAGC	849
Db	1639	 CAGGAGCCACAGCAGCGGGAGCCACAGCAGCGGAGGCCCCCAGCAGCGGGAGCCACAGCAG	1698
Qy	850	TTTGGAGAGAAGAGCTTGCCTCCAGCCTGGCCTACGAGAGCGCGCGCGCTGCAGG	909
Db	1699	 CGGAGCCACAGCAGCGGGAGCCACAGCAGCGGAGCCACAGCAGCGGGAGCCACAGCAG	1758
Qy	910	GACGAGCTGGAGGCGCGGAGCCCAAGCTCAAGCAGGCGCTCCGACAGG	969
Db	1759	CGGAGCCACAGCAGCGGGAGCCACAGCAGCGATGAGCAGCAGGATGAGCAGCAG	1818
Qy	970	AGCCAGCGCGCAGCAGAGCTCTGCACTTGCAGGTTACTGCAGCTTCAGCAGGAGAGCGG	1029
Db	1819	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCGGATGAGCAGCAGCAGGATGAGCAGCAG	1878
Qy	1030	CAGCTCCGCGCAGGAGCTCGAGAGCCTCATGAAGSAGCAGCAGCTGCTGGAGAGCCAAAGCTC	1089
Db	1879	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCGATGAGCAGCAGCAGGATGAGCAGCAG	1938
Qy	1090	AGGTCTTACGAGAGGGAGAGACACAGCTTCGGCCCCCGCTGGAGGAGAGCCCAAGTGGGAG	1149
Db	1939	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCGGATGAGCAGCAGCAGGATGAGCAGCAG	1998
Qy	1150	GTGTGCCAGAACTCAGCGAGATCTCCCTCTCTGAAGCAGCAGCTGGAAGGAGTCCCAGACG	1209
Db	1999	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCGATGAGCAGCAGCAGGATGAGCAGCAG	2058
Qy	1210	GAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCAACAGCTGGAAGGACACGCGG	1269
Db	2059	CAGGATGAGCAGCAGCAGGATGAGCAGCAGCGGATGAGCAGCAGCAGGATGAGCAGCAG	2118
Qy	1270	GGCAAGCTGGAGGCGCTGGAGCTGAGAGCCAGGACCTGGAGGCGCCCTCGCGCACCAAG	1329
Db	2119	GAGCAGCAGGATGAGCAGGAGCAGCAGATGAGCAGCAGCAGGATGAGCAGCAGCAGCAG	2178
Qy	1330	GGCCTGGAGCTGGAGGCTCTGTGGAATGAGCTGCAGCGCAAGAGAACGAGGCGGAGCTG	1389
Db	2179	GATCAGCAGCAGCAGCAGGATGAGCAGCAGCGATGAGCAGCAGCAGCAGCAGGATGAG	2238
Qy	1390	CTCGGGAGAGGTGAACCTGCTGGAGCAGGAGCTGCAGAGCTGCGGGCCCCAGGCCGCC	1449
Db	2239	CAGCAGCAGCAGGATGAACAGGAGCAGCAGGAGCAGCAGGAGCAGCAGGAGGAGCAGGAG	2298
Qy	1450	CTGCCCCCGCATATGGGCGCCCACTTCCCCAGGACGCTCCCTCGCCCTGCAGCGGGAG	1509
Db	2299	CAGGATTAGAGGAGCAGGAGCAGGATTAGAGGATCAGGAGCAGGAGTTAGA--GGAG	2355
Qy	1510	CTGAGCGGCTGCGGGCGGAGCTCGGGAGGAGCGGCAGGCCATGACAGATGTCTCTCG	1569
Db	2356	CAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAG	2415
Qy	1570	GGCTTCCAGCATGACGGCTCGTGTGAAGGAGGAGAGAGGATGATTTCAGTACCAAG	1629
Db	2416	GAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGATTAGAGGAGCAGGAGCAGGAGTTA	2475
Qy	1630	AAACAGCTGCAGAGAGCTACGTGGCCATGTACAGCGGAAACACAGCCCTGGAGAGGCC	1689
Db	2476	GAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGCAGGAGTTAGAGGAGCAGGAGGTCGAA	2535
Qy	1690	CTGCAGCAGCTGGCACGTGGGACACAGCGCCGGGAGCCCTTGGAGGTTGACCTCGAAAGG	1749
Db	2536	GAGCAGGAGCAGGAGGTGGAAGACAGCAGCAGGAGCAGGAGCAGGAGGATTAGAGGAG	2595
Qy	1750	GCTGACATCCCTTAGAGGA	1769
Db	2596	GTGGAGGAGCAAGAGCAGGA	2615

RESULT 14

US-09-298-568-1

US-09-298-368-1  
; Sequence 1, Application US/09298568





SUMMARIES									
Result No.	Query			DB	ID	Description			
	Score	Match	Length						
1	808.8	45.2	2411	9	US-09-764-904-89	Sequence 89, Appl			
2	808.8	45.2	2411	9	US-10-091-548-89	Sequence 89, Appl			
3	808.8	45.2	2411	9	US-10-074-095-1133	Sequence 1133, Ap			
4	808.8	45.2	2411	10	US-09-764-860-1133	Sequence 1133, Ap			
5	560	31.3	560	9	US-09-764-904-17	Sequence 17, Appl			
6	560	31.3	560	9	US-10-091-548-17	Sequence 17, Appl			
7	560	31.3	560	9	US-10-074-095-269	Sequence 269, App			
8	560	31.3	560	10	US-09-764-860-269	Sequence 269, App			
9	314	17.5	5257	9	US-10-171-581-63	Sequence 63, Appl			
10	258.2	14.4	1617	10	US-09-925-300-60	Sequence 60, Appl			
11	131.2	7.3	2108	10	US-09-862-832-225	Sequence 225, App			
12	114.6	6.4	1852	10	US-09-969-852-4	Sequence 4, Appl			
13	114	6.4	561	10	US-09-917-800A-194	Sequence 194, App			
14	107.2	6.0	14800	10	US-09-954-456-1601	Sequence 1601, Ap			
15	105.2	5.9	22585	9	US-09-764-891-6987	Sequence 6987, Ap			
16	89	5.0	7065	9	US-08-991-456-115	Sequence 115, App			
17	89	5.0	7065	10	US-09-874-523-115	Sequence 115, App			
18	87.2	4.9	3773	10	US-09-925-302-47	Sequence 47, Appl			
19	87.2	4.9	6457	10	US-09-880-107-3389	Sequence 3389, Ap			

QY 645 GGGCATCGTCTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 704  
Db 1606 GGGCATCGTCTCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCTTCTCCGACGG 1665  
QY 705 AGGTAGCAAGCTGGGCACTCGAAACAAGGACAGCAAGGGCCCTCGTGTGTCGGCTCCCC 764  
Db 1666 AGGTAGCAAGCTGGGCACTCGAAACAAGGACAGCAAGGGCCCTCGTGTGTCGGCTCCCC 1725  
QY 765 CATCTCCAGGACGAGTGCAGATCCAGGAGCTGGAGCAGAGAGCTGTGGAGAGGAGGG 824  
Db 1726 CATCTCCAGGACGAGTGCAGATCCAGGAGCTGGAGCAGAGAGCTGTGGAGAGGAGGG 1785  
QY 825 CGCCCTCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCCTA 884  
Db 1786 CGCCCTCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCCTA 1845  
QY 885 CGAGGAGCGCGCGCGCTGTCAGGAGCAGCTGGAGGGCCCGGAGCCCAAGGCGGCAA 944  
Db 1846 CGAGGAGCGCGCGCGCTGTCAGGAGCAGCTGGAGGGCCCGGAGCCCAAGGCGGCAA 1905  
QY 945 CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCAGAGGCTCTGCACTGCGAGGT 1004  
Db 1906 CAAGCTCAAGCAGGCTCGAGAGAGCCAGCGCGCAGAGGCTCTGCACTGCGAGGT 1965  
QY 1005 ACTGCAGCTTCAGCAGGAGAGCGGCGAGCTCCGGCAGGAGCTTCGAGAGCTCATGAAGGA 1064  
Db 1966 ACTGCAGCTTCAGCAGGAGAGCGGCGAGCTCCGGCAGGAGCTTCGAGAGCTCATGAAGGA 2025  
QY 1065 GCAGGACCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGGGCC 1124  
Db 2026 GCAGGACCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGGGCC 2085  
QY 1125 CGCGCTGAGGAGACCAAGCTGGAGGTGTGCC 1156  
Db 2086 CGCGCTGAGGAGACCAAGCTGGAGGTGTGCC 2117

## RESULT 2

US-10-091-548-89  
; Sequence 89, Application US/10091548  
; Publication No. US20030049703A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA122CI  
; CURRENT APPLICATION NUMBER: US/10/091,548  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 137  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 89  
; LENGTH: 2411  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-091-548-89

Query Match. 45.2%; Score 808.8; DB 9; Length 2411;  
Best Local Similarity 99.8%; Pred. No. 5.5e-178;  
Matches 810; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 345 GGGCTCCGAGAGGCTGAGTGAGGCCACACAGCCTTCAAGCTGTGCTGCCACGCTCAGG 404  
Db 1306 GGGCTCCGAGAGGCTGAGTGAGGCCACACAGCCTTCAAGCTGTGCTGCCACGCTCAGG 1365  
QY 405 AGCCATCTGCACTCTCTCCCGAGAGTGCCAGCCAGCAGTGCACCCGCCCTCCAGA 464  
Db 1366 AGCCATCTGCACTCTCTCCCGAGAGTGCCAGCCAGCAGTGCACCCGCCCTCCAGA 1425  
QY 465 CAGGCCCCAGGAGGAGGAGCTGAAGCTTGGCTGTGGGGCGCTGTGAGACTCCGG 524  
Db 1426 CAGGCCCCAGGAGGAGGAGCTGAAGCTTGGCTGTGGGGCGCTGTGAGACTCCGG 1485

QY 525 CCGGAATCTCATGTCCAGCCTGCCACACACAGCACCAGCAGCAGCTACAGCTGGACCC 584  
Db 1486 CCGGAATCTCATGTCCAGCCTGCCACACACAGCACCAGCAGCAGCTACAGCTGGACCC 1545  
QY 585 GTGGTCAACCCGTGGGACCCACAGCCGTTTGGGGGTTCGCCCCACAACTCACCAC 644  
Db 1546 GTGGTCAACCCGTGGGACCCACAGCCGTTTGGGGGTTCGCCCCACAACTCACCAC 1605  
QY 645 GGCATCGTCTCCAGACAGCAACATGATGAGCCTGAAGGCTGTGCTTCTCCGACGG 704  
Db 1606 GGCATCGTCTCCAGACAGCAACATGATGAGCCTGAAGGCTGTGCTTCTCCGACGG 1665  
QY 705 AGGTAGCAAGCTGGGCACTCGAAACAAGGACAGCAAGGGCCCTCGTGTGTCGGCTCCCC 764  
Db 1666 AGGTAGCAAGCTGGGCACTCGAAACAAGGACAGCAAGGGCCCTCGTGTGTCGGCTCCCC 1725  
QY 765 CATCTCCAGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTGGAGAGGAGGG 824  
Db 1726 CATCTCCAGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTGTGGAGAGGAGGG 1785  
QY 825 CGCCCTCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCCTA 884  
Db 1786 CGCCCTCAGAGCTGCAGCGCAGCTTTGAGGAGAGGAGCTTGCCTCCAGCCTGGCCTA 1845  
QY 885 CGAGGAGCGCGCGCGCTGTCAGGAGCAGCTGGAGGGCCCGGAGCCCAAGGCGGCAA 944  
Db 1846 CGAGGAGCGCGCGCGCTGTCAGGAGCAGCTGGAGGGCCCGGAGCCCAAGGCGGCAA 1905  
QY 945 CAAGCTCAAGCAGGCTTCGCAGAGAGCCAGCGCGCAGAGTCTCTGCACTTGCAGGT 1004  
Db 1906 CAAGCTCAAGCAGGCTTCGCAGAGAGCCAGCGCGCAGAGTCTCTGCACTTGCAGGT 1965  
QY 1005 ACTGCAGCTTCAGCAGGAGAGCGGCGAGCTCCGGCAGGAGCTTCGAGAGCTCATGAAGGA 1064  
Db 1966 ACTGCAGCTTCAGCAGGAGAGCGGCGAGCTCCGGCAGGAGCTTCGAGAGCTCATGAAGGA 2025  
QY 1065 GCAGGACCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGGGCC 1124  
Db 2026 GCAGGACCTGTGGAGACCAAGCTCAGGTCTTACGAGAGGAGAGACCAAGCTTCGGGCC 2085  
QY 1125 CGCGCTGAGGAGACCAAGCTGGAGGTGTGCC 1156  
Db 2086 CGCGCTGAGGAGACCAAGCTGGAGGTGTGCC 2117

## RESULT 3

US-10-074-095-1133  
; Sequence 1133, Application US/10074095  
; Publication No. US2003007704A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008CI  
; CURRENT APPLICATION NUMBER: US/10/074,095  
; CURRENT FILING DATE: 2002-02-14  
; PRIOR APPLICATION NUMBER: 09/764,860  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14



; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 60/225,266  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/249,218  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,208  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,213  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,212  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,207  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,245  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,244  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,217  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,211  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,215  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,264  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,214  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,297  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/232,400  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: 60/231,242  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/232,081  
; PRIOR FILING DATE: 2000-09-08





Oy 851 TTGAGGACAGGAGCTTCCC 870  
Db 541 TTGAGGACAGGAGCTTCCC 560

## RESULT 7

US-10-074-095-269  
; Sequence 269, Application US/10074095  
; Publication No. US2003007704A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008C1  
; CURRENT APPLICATION NUMBER: US/10/074,095  
; CURRENT FILING DATE: 2002-02-14  
; PRIOR APPLICATION NUMBER: 09/764,860  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/225,757  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/226,868  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/216,647  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,267  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/216,880  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: 60/225,270  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/251,869  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/235,834  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/234,274  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/234,223  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: 60/228,924  
; PRIOR FILING DATE: 2000-08-30  
; PRIOR APPLICATION NUMBER: 60/224,518  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,369  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/224,519  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,964  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/241,809  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/249,299  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/236,327  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/241,785

; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/244,617  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 60/225,268  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/236,368  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/251,856  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/251,868  
; PRIOR FILING DATE: 2000-12-08  
; PRIOR APPLICATION NUMBER: 60/229,344  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/234,997  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: 60/229,343  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,345  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,287  
; PRIOR FILING DATE: 2000-09-01  
; PRIOR APPLICATION NUMBER: 60/229,513  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/231,413  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: 60/229,509  
; PRIOR FILING DATE: 2000-09-05  
; PRIOR APPLICATION NUMBER: 60/236,367  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/237,039  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,038  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/236,370  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: 60/236,802  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,037  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/237,040  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: 60/240,960  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/239,935  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/239,937  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 60/241,787  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/246,474  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/246,532  
; PRIOR FILING DATE: 2000-11-08  
; PRIOR APPLICATION NUMBER: 60/249,216  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/249,210  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/226,681  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,759  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,213  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/227,182  
; PRIOR FILING DATE: 2000-08-22  
; PRIOR APPLICATION NUMBER: 60/225,214  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/235,836  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/230,438  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/215,135  
; PRIOR FILING DATE: 2000-06-30

; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 31.3%; Score 560; DB 9; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGGTGCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGGTGCAGTGAGGC 60
Qy 371 CCACAGCCTTCAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTGTCACCTCTCCCGGAGA 430

Db 61 CCACAGCCTTCAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTGTCACCTCTCCCGGAGA 120
Qy 431 GTGCCAGCCACAGCTGCACCCGCCCTCCAGACAAGCCCCCAAGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACAGCTGCACCCGCCCTCCAGACAAGCCCCCAAGAGCAGAGCTGAAGC 180
Qy 491 CTGGCCTGTGCTCTGGGGCGCTGTGCAGACTCCGGCGGAACCTCCATGTCCAGCCTGCCCA 550
Db 181 CTGGCCTGTGCTCTGGGGCGCTGTGCAGACTCCGGCGGAACCTCCATGTCCAGCCTGCCCA 240
Qy 551 CACACAGCACCAGCAGCAGCTACAGCTGGACCGCTGGTCCACACCCGTGGAGCCCAAA 610
Db 241 CACACAGCACCAGCAGCAGCTACAGCTGGACCGCTGGTCCACACCCGTGGAGCCCAAA 300
Qy 611 GCCGTTTGGGGGCTCCGCCCCACAAATCACCAGGCGCATCTCTCCAGACACAGCAACA 670
Db 301 GCCGTTTGGGGGCTCCGCCCCACAAATCACCAGGCGCATCTCTCCAGACACAGCAACA 360
Qy 671 TGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCGCACTCGAACA 730
Db 361 TGATGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCGCACTCGAACA 420
Qy 731 AGGCAGACAAAGGGCCCTCGTGTGCTGCCCTCCCATCTCCACGACAGTGCAGCATCC 790
Db 421 AGGCAGACAAAGGGCCCTCGTGTGCTGCCCTCCCATCTCCACGACAGTGCAGCATCC 480
Qy 791 AGGAGCTGGAGCAGAAAGCTGTTGGAGAGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCT 850
Db 481 AGGAGCTGGAGCAGAAAGCTGTTGGAGAGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCT 540
Qy 851 TTGAGGAGAGGAGCTTGCC 870
Db 541 TTGAGGAGAGGAGCTTGCC 560

RESULT 8
US-09-764-860-269
; Sequence 269, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 269
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-269

Query Match 31.3%; Score 560; DB 10; Length 560;
Best Local Similarity 100.0%; Pred. No. 1.7e-120;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGGTGCAGTGAGGC 370
Db 1 CCAAGCTCATGCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAAAGGGTGCAGTGAGGC 60
Qy 371 CCACAGCCTTCAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTGTCACCTCTCCCGGAGA 430
Db 61 CCACAGCCTTCAGCCTGTGTGCTGCCAGCGTCAGGAGCCATCTGTCACCTCTCCCGGAGA 120
Qy 431 GTGCCAGCCACAGCTGCACCCGCCCTCCAGACAAGCCCCCAAGAGCAGAGCTGAAGC 490
Db 121 GTGCCAGCCACAGCTGCACCCGCCCTCCAGACAAGCCCCCAAGAGCAGAGCTGAAGC 180
Qy 491 CTGGCCTGTGCTCTGGGGCGCTGTGCAGACTCCGGCGGAACCTCCATGTGTCAGCCTGCCCA 550
Db 181 CTGGCCTGTGCTCTGGGGCGCTGTGCAGACTCCGGCGGAACCTCCATGTGTCAGCCTGCCCA 240

Qy	551	CACACAGCACGACGAGCTACAGCTGGAACCGCTGGTGTCAACCCGTTGGGACCCACAA	610
Db	241	CACACAGCACGACGAGCTACAGCTGGAACCGCTGGTGTCAACCCGTTGGGACCCACAA	300
Qy	611	GCGTTTTTGGGGGCTCCGCCCAACAATACCCAGGGCATCTGCTCCAGGACAGCAACA	670
Db	301	GCGTTTTTGGGGGCTCCGCCCAACAATACCCAGGGCATCTGCTCCAGGACAGCAACA	360
Qy	671	TGATGACCTTGAGGCTCTGCTTCTCCAGCGGAGTAGCAAGCTGGGGCACTCGAACA	730
Db	361	TGATGACCTTGAGGCTCTGCTTCTCCAGCGGAGTAGCAAGCTGGGGCACTCGAACA	420
Qy	731	AGGCAGACAAGGGCCCTCGTGTGTCGCTCCCCCATCTCCAGCGAGGTGAGCATCC	790
Db	421	AGGCAGACAAGGGCCCTCGTGTGTCGCTCCCCCATCTCCAGCGAGGTGAGCATCC	480
Qy	791	AGGAGCTTGGAGCAGAAAGCTGTTGGAGAGGAGGGCGCCCTCCAGAAGCTGCAGCGAGCT	850
Db	481	AGGAGCTTGGAGCAGAAAGCTGTTGGAGAGGAGGGCGCCCTCCAGAAGCTGCAGCGAGCT	540
Qy	851	TTGAGGAGAAGGAGCTTTGCC	870
Db	541	TTGAGGAGAAGGAGCTTTGCC	560

## RESULT 9

```

US-10-171-581-63
; Sequence 63, Application US/10171581
; Publication No. US2003010426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 63
; LENGTH: 5257
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: AB011124
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-63

```

	Query Match	17.5%	Score 314;	DB 9;	Length 5257;
	Best Local Similarity	58.2%;	Pred. No. 1.8e-63;		
	Matches 619;	Conservative 0;	Mismatches 400;	Indels 45;	Gaps 2;
Qy	710	GCAAGCTGGGGCACTCGAAACAAGGCACAAGGGCCCTCTGTGTGTCGGCTCCGCCATCT	769		
Db	2292	GCCACCTGGGCTCTGGGAGGGCGAGGTGGAGGCGCTTTCGGCGGCTGTCTACCGC	2351		
Qy	770	CCACGGACGAGTGCAGCATCCAGAGCTGGAGCAGAAGACTGTGGAGAGGAGGGCGCCC	829		
Db	2352	CCTCCCCAGTGCACCTCATCCAGAGCTGGAGGAGCGGCTGTGGGAAAGAGCAGGAGG	2411		
Qy	830	TCCAGAAGCTGCACGCGCAGCTTTTCAGGAGAGGAGCTTGGCTCCAGCGCTTGCCTTACGAGG	889		
Db	2412	TGGCAGCTCTGCGCGCAGCGCTGTGAGCAGAGCGGCGGTGTGGCCCAAGGTACTGGAGG	2471		
Qy	890	AGCGGCGCGCGCTGCAGGACAGCTGGAGGCCCGCGAGCCCAAGGCGGCAACAGC	949		
Db	2472	AGCGGCAGAACGCTGGAGCGGGAGCTTGGCCGAGCTGCGGCAGGGCTGCAGCGGGAAAGC	2531		
Qy	950	TCAAGCAGGCCCTCCGAGAAGAGCCAGCGCGCAGCAGGTTCTTGCACCTTCGAGGTACTGC	1009		
Db	2532	TACAGCAGTGGCCCGCAGCTGCCCAGCGCCCGCAGCAGGGGCTACAGTGCAGGTGTTC	2591		

Qy	1010	AGCTTCAGCAGGAGAACGGCGCAGCTCCGGCAGAGAGCTCAGAGACCTCATGAAGCAGCAGG	1069
Db	2592	GGCTGCAGCAGGACAAGAACAGCAGCTCAGAGAGAGAGCGGCCGGCTGATCGCGCAGCGGG	2651
Qy	1070	ACCTGCTGGAGACCAAGCTCAGTCTCTACGAGAGGGAGAGACCAGCTTCGGGCCCGCGC	1129
Db	2652	AAGAGCTGGAGGACAAGGTGGCCGCTGCGCAGAAGGAGCAGGCGCGACTTCCTGCCCCCGGA	2711
Qy	1130	TGAGGAGAGACCCAGTGGGAGGTGTGCAGAAAGTCAGCGAGATCTCCCTCCTGAAGCAGC	1189
Db	2712	TAGAGGAACTAAGTGGAGGTGTGCAGAAAGCTGGCAGATCTCCCTCCTGAAGCAGC	2771
Qy	1190	AGCTGAAGGAGTCCCAGACGAGGTGAACGCCAAGGCTAGCGAGATCTCTGGGTCTCAAGG	1249
Db	2772	AGCTGAAGGACTCGCAGGCGGATGTGTGCAGAAAGTTGAGTGAGATCGTGGGACTGCGCT	2831
Qy	1250	CACAGCTGAAGGACACGCGGGGCAAGCTGAGAGGGCCTGGAGCTGAGGACCCAGGACCTGG	1309
Db	2832	CGCAGCTGCGGAGGGCCGGGCTTCGCTCGGAGAAAGGAGGACAGCTGCTCAGCCTGC	2891
Qy	1310	AGGCGGCCCTGCGCACCAAGGGCTCGAGCTGAGGTCTGTGAGAAATGAGCTGCAGCGCA	1369
Db	2892	GGGACTCCTTCAGCAGCAGCAGCCAGCCTGAGCTGGCGAAGGCGAGCTGCCCTGCCG	2951
Qy	1370	AGAAGAACAGGCGGAGCTGCTCGGGAGAAAGGTGAACCTGTGGAGCAGGAGCT-----	1424
Db	2952	CCTGCCTCAAGCGGGGCTGACCCCGCTGACCCGCGCAGCCGAGCAGGATGCTCTGGCCA	3011
Qy	1425	-----GCAGGAGCTCGGGCCCGCAGCGCCCTCGGCCCGCGACATGGGC	1468
Db	3012	CCTGCGAGACGACGAGGCTAAGATGCCCGTCAGGCGGGGTGGCCGCTGCGCGCTCCT	3071
Qy	1469	CGGCCACTTCCCGCAGGACG-----TCCCTGCCCTTCGAGC	1504
Db	3072	TGGTTTCCTGAGCGGGAGCGGAGGCTGGCGGGAGAGCGGGAACGGGGCCTTCGCGC	3131
Qy	1505	GGGAGCTGGAGCGGCTGCGGGCCGAGCTGCGGAGAGCGGCAAGGCCATGACCAAGATG	1564
Db	3132	GGGAGGTGGGGCGGCTGCAGGCCGAGCTGCGCGCTGAGCGCGGGGCCGGGAGCGCAGG	3191
Qy	1565	CCTCGGGCTTCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAGAGAGGTGATTCACT	1624
Db	3192	GTGCCAGCTTCGCGGAGGACGCGCGGTGTGGCTGGAGGAGAGGAGAGGTGATCGAGT	3251
Qy	1625	ACCAGAAACAGCTGCACACAGCTACTGTGCCCATGTACACGCGGAACCAAGCGCTGGAGA	1684
Db	3252	ACCAAGACAGCTGCAGCTGAGCTACGTGAGATGTACACGCGCAACCAAGCAGCTGGAGC	3311
Qy	1685	AGGCCCTGCACGACCTGGCAGCTGGGACAGCGCCCGGGAGCC	1728
Db	3312	GCAGGCTGCGGGAGCGCGGGGCCGACGGGGGTGCAAGCAGCC	3355

RESULT 10  
US-09-925-300-60  
; Sequence 60, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925.300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 60  
; LENGTH: 1617

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1590)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1592)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1595)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1617)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-60

Query Match      14.4%; Score 258.2; DB 10; Length 1617;
Best Local Similarity 60.2%; Pred. No. 1.3e-50;
Matches 523; Conservative 0; Mismatches 303; Indels 43; Gaps 4;

Qy 962 CGCAGAAGACGAGCGCGCAGCAGGCTCTGCACCTGCAGTACTGCAGCTTTCAGCAGG 1021
Db 33 CACACGGGCGACAGCGGCGCCCAAGWTGTCAGCTGCAGTGTTCAGTGTCA-CAGG 91

Qy 1022 AGAAGCGGCGAGCTCCGGCAGGAGCTCGAGAGCTTCATGAAGGAGCAGGACCTGTCTGGAGA 1081
Db 92 AGAAGCGCAATTGCGAGGACGACTTCGCACAGCTGCTGCAGGAGCGGCAACAGCTGGAGC 151

Qy 1082 CCAAGCTCAGTCTTACAGAGGGGAGAGACAGCTTCGGCCCGCCGCTGCAGGAGACCC 1141
Db 152 GCGCTGCGCCACCTTTGAGGCGGGACAGCGGGAGC-TCGGGCGGAGGCTTGAGGAGACCA 210

Qy 1142 AGTGGAGGTGTGCCAGAGTTCAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAGT 1201
Db 211 AGTGGAGGTGTGCCAGAAATCAGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGGAGT 270

Qy 1202 CCCAGA CCGAGGTGAA CCGCAAGGCTAGCGAGATCTCTGGGTCTCAAGGCA CAGCTGAAAG 1261
Db 271 CTCAGGCAGAGCTGTGTGCAGAAGGCGCAGCTGTGGTCTCGGGGTGCGCTCGCGG 330

Qy 1262 ACACGCGGGCAAGCTGAGGCGCTGAGCTGAGGACCCAGGACCTGAGGCGCCCTGC 1321
Db 331 AGGCGCGTGTACGCTGCGGGTCAAGTGTAGGCGCGGTGCGCGGGTCTACAGAGGCGCGCC 390

Qy 1322 GCACCAAGGCTTGAGCTGAGGTCTGTGAGATGAGCTGCAGCGCAAGAGAACAGG 1381
Db 391 GAGCTCGGAGCTGAGCTGAGGCTGTTCAGGAGCTGCAGGCAACCGCCAGGAAG 450

Qy 1382 CGGAGCTGCTCGGGAGAGGTGAACTGCTGGAGCAGGAGCTGCAGGAGCTGCGGG--- 1438
Db 451 CTGAGCAGCTGCGGGAGAAAGCTGGGCAGTTGGATGCTGAGCGCGCGGACTCCGGGAGC 510

Qy 1439 -----CCGAGCGCCCTGCGCCCGGCAATGG 1465
Db 511 CCCCTGTGCCACCTGCCACCGCTGACCCATTCCTCTGGCAGAGAGTATAGGCCAAAG 570

Qy 1466 GCGCGCCACCTTCCCGCAGACGTCCTGCGCTGCAGCGGAGCTGAGCGGCTGCGGG 1525
Db 571 TGCAGCGGCGAGCAGCCGGGTGTTGGGCGAGCTTGCGGGCCAGGTGAGGCGATTGCGGG 630

Qy 1526 CCGAGCTCGGGAGGAGCGCAAGGCCATGACCCAGATGCTCTCGGGCTTCAGCATGAGC 1585
Db 631 TGGAGCTCAGCGGAGCGCGCGGGGTGAGGAGCAGCGGACAGCTTTGAGGGGAGC 690

Qy 1586 GGCTGCTGGAAGAGGAGAGAGAGAGTGTATTCACTACAGAAAACAGCTGCAGCAGA 1645
Db 691 GGTGGCTGCGGCGCAGAGAGGAGCAGGTGATCCGCTACAGAAAGCAGCTGCAGCACA 750

Qy 1646 GCTAGTGGCCATGTACAGCGGACAGCGCTTGGAAAGCCCTGCAGCAGCTGGCAC 1705
Db 751 ACTACATCCAGATGTACCGGCGCAACCGCGAGCTAGACAGAGAGCTGCAGCAGCTCAGCC 810
```

```
Qy 1706 GTGGGACAGCGCCGGGAGGCCCTTGGAGGTGACCTG-----GAAGGGGTGACATCCC 1760
Db 811 TGGAGCTGGAGCGCCGGGAGCTCGCTGACCTGGGCTTGGCCGAGCAGCCCTGCTGCTG 870

Qy 1761 CTACGAGACATCATAGCCACTGAGATCT 1789
Db 871 CCTGGAGGAGATCACTGCTACTGAGATCT 899

RESULT 11
US-09-962-832-225
; Sequence 225, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 225
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-962-832-225

Query Match      7.3%; Score 131.2; DB 10; Length 2108;
Best Local Similarity 48.4%; Pred. No. 3.3e-21;
Matches 429; Conservative 0; Mismatches 448; Indels 9; Gaps 2;

Qy 860 AGGAGCTTGCTCCAGCTTCAGGAGCGCGCGCGCTGCGAGGACGAGCTGCAGAGCGG 919
Db 499 AGCAGCAGGAGGGGCACTGAAGCACCTAGAGCAGCAGGAGGAGCAGCTGAAGCACCCGG 558

Qy 920 AGGCGCGGAGCCCAAGGGCGCAACAGCTCAAGCAGGCGCTCCGAGAGAGCGAGCGG 979
Db 559 AGCAGCAGGAGGGGCGAGCTGGAGCTCCAGAGCAGCAGGAGGGCGAGCTGAGCTCCAG 618

Qy 980 CGCAGCAGGCTCTCCACTGCAGCTACTGCACTTTCAGCAGGAGAGCGGCGAGCTCCGGC 1039
Db 619 AGCAGCAGGAGGGGCGAGCTGAGCTCCAGAGCAGCAGGAGGGGCGAGCTGAGCTCCAG 678

Qy 1040 AGGAGCTCGAGAGCTCATGAAGGAGCAGGACCTGCTGGAGACCAAGCTCAGGCTCCTACG 1099
Db 679 AGCAGCAGGAGGGGCGAGCTGGAGCTCCACAGCAGCAGGAGGGGCGAGCTGAGCTCTCTG 738

Qy 1100 AGAGGAGAGACACAGCTTGGCGCCCGGCTGAGGAGAGACCCAGTGGAGAGGTGTGCCAGA 1159
Db 739 AGCAGCAGGAGGGGCGAGCTGGAGCTCTGAGCAGCAGGAGGGGCGAGCTGAGCTCTCTG 798

Qy 1160 AGTCAGGCGAGATCTCCCTCTGAAGCAGCAGCTGAAGAGCTCCAGACGCGAGGTGAACG 1219
Db 799 AGCAGCAGGAGGGGAGCAGCTGAAGCACTTGGAGCAGCAGGAGGGGCGAGCTGAGGCTCCAG 858

Qy 1220 CCAAGGCTAGCGAGATCTTGGGTCTCAAGGCACAGCTGAAGGAGACACGCGGGCGCAAGCTGG 1279
Db 859 AGGAGCAGATGGGCGAGCTGAAGTACCTGGAACAGCAGGAGGGGCGAGCTGAAGCACCTGG 918

Qy 1280 AGGCGCTGGAGCTGAGGAC-----CAGGACCTGAGGAGGGCGCCCTTGGCGACCAAGGGGCC 1333
Db 919 ATCAGCAGGAGAGCAGCAGCAGAGCTCCAGAGCAGCAGATGGGGCAGCTGAAGCACCTGG 978

Qy 1334 TGGAGCTGGAGGTCTGTGAGATGAGCTGCAAGCCCAAGAAACAGCGCGGAGCTGCTGC 1393
Db 979 AGCAGCAGGAGGGGCGAGCTTAAGCATCTGGAGCAGCAGGAGGGGCAACTGGAGCAGCTGG 1038

Qy 1394 GGGAGAAAGGTGAACCTCTGCTGGAGCAGGAGCTGCAGGAGCTCGGGGCCCGAGCTCGCCCTGG 1453
```





; PRIOR APPLICATION NUMBER: US 60/295,798  
 ; PRIOR FILING DATE: 2001-06-06  
 ; PRIOR APPLICATION NUMBER: US 60/297,457  
 ; PRIOR FILING DATE: 2001-06-13  
 ; PRIOR APPLICATION NUMBER: US 60/298,884  
 ; PRIOR FILING DATE: 2001-06-19  
 ; PRIOR APPLICATION NUMBER: US 60/303,459  
 ; PRIOR FILING DATE: 2001-07-09  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 194  
 ; LENGTH: 561  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA875041  
 US-09-917-800A-194

Query Match 6.4%; Score 114; DB 10; Length 561;  
 Best Local Similarity 77.5%; Pred. No. 2.7e-17;  
 Matches 138; Conservative 0; Mismatches 40; Indels 0; Gaps 0;  
 QY 772 ACGGACGAGTCAGCATCCAGGAGCTGGAGAGAGAGCTTTGCGAGAGGGCGCCCTC 831  
 Db 561 ACGGATGAGTCACCATTCAGAGCTGGAGAGAGAGCTTGCTCGGAGAGCTGCACTA 502  
 QY 832 CAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTGCTCCAGCTGGCCTTACGAGGAG 891  
 Db 501 CAGAGCTACAGCGCAGTTTCGATGAGAGAGAGTTTGCTTCGCGCAGCTTCGAGGAG 442  
 QY 892 CGGCGCGCGCTGCAGGAGAGCTGGAGGCGCGGAGCGCCCAAGGCGGCAACAGC 949  
 Db 441 CGGCGACGCGGACTAGAGAGCAACTGGAGTGCTGGAACCTAGAGCAAGCTGAAGC 384

# RESULT 14

US-09-954-456-1601  
 ; Sequence 1601, Application US/09954456  
 ; Patent No. US20020115057A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
 ; TITLE OF INVENTION: Sets  
 ; FILE REFERENCE: 689290-76  
 ; CURRENT APPLICATION NUMBER: US/09/954,456  
 ; PRIOR FILING DATE: 2001-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/233,617  
 ; PRIOR FILING DATE: 2000-09-18  
 ; PRIOR APPLICATION NUMBER: US/60/234,052  
 ; PRIOR FILING DATE: 2000-09-20  
 ; PRIOR APPLICATION NUMBER: US/60/234,923  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,134  
 ; PRIOR FILING DATE: 2000-09-25  
 ; PRIOR APPLICATION NUMBER: US/60/235,637  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,638  
 ; PRIOR FILING DATE: 2000-09-26  
 ; PRIOR APPLICATION NUMBER: US/60/235,711  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,720  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,840  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: US/60/235,863  
 ; PRIOR FILING DATE: 2000-09-27  
 ; NUMBER OF SEQ ID NOS: 2276  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1601  
 ; LENGTH: 14800  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-954-456-1601

Query Match 6.0%; Score 107.2; DB 10; Length 14800;  
 Best Local Similarity 46.1%; Pred. No. 1.4e-15;  
 Matches 447; Conservative 0; Mismatches 508; Indels 15; Gaps 2;  
 QY 791 AGGAGCTGGAGCAGAAAGCTGTGGAGAGGGAGGCGCCCTCCAGAACTGTCAGCGCAGCT 850  
 Db 5873 AGGACACGCTGAGCAGCGCGGCGAGGTGGAGGAAGAGATCTGCGCTGAAGCGGAGCT 5932  
 QY 851 TTGAGGAGAGGAGCTTGCTCCAGCTCCAGCTGCGCTACGAGGAGGCGCGCGCTCCAGGG 910  
 Db 5933 TCGAGAAAGCGCGCTGCGCAAGCGGAGCTGGAGCTGGAGCGCATCCGCACGA 5992  
 QY 911 ACGAGCTGGAGGGCGCGGAGCGCCAAAGCGCGCAACAGCTCAAGCAGGACCTCGCAGAAAGA 970  
 Db 5993 ACGCGGAGGACACCTGCGCAGCAAGGAGCAGGCGGAGCTGGAGCGCGGAGGACCGCG 6052  
 QY 971 GCCAGCGCGCGCAGAGGTCTTCACCTGTCAGGTACTGTCAGCTTCAGCAGGAGAAAGCGCG 1030  
 Db 6053 AGCTGGCGCGCGAGGAGGCGCGCGCTGAGGCTGAGGAGCGCTGCGAGAGAGCG 6112  
 QY 1031 AGCTCCGCGCAGGAGCTCGAGAGCTCATGAAGGAGCAGGACCTGCTGTCAGAGACCTCA 1090  
 Db 6113 TGGCGCGCGAGGAGGAGCGCGCACCGGACGCGAAGGCGCGCTGGAGGAATCGAGCGCG 6172  
 QY 1091 GGTCTCTACGAGAGGAGAAAGACACAGCTTCGGCGCGCGCTGGAGGAGACCCAGTGGGAGG 1150  
 Db 6173 TGAAGCCACGCTGGAGGAGCGCGCGCTGCGGAGCGAGCGGAGCAGAGTCCGCGC 6232  
 QY 1151 TGTCCAGAAAGTCAAGCGAGATCTCTCTGTAAGCAGCAGCTGAAGAGTCCCAAGCGG 1210  
 Db 6233 GGCAGCTGCAGCTGCGCGCGAGGAGCGCGCCAGAGCGGCTG-----CAGGCGG 6280  
 QY 1211 AGGTGAACGCGCAAGGCTAGCGAGATCTTGGGTCTCAAGGCGACAGCTGAAGGACAGCGCGG 1270  
 Db 6281 AAGAGAAAGGACACGCTTCGCGGTGCGAGAGAGGAGCAGAGCTACAGCAGACGCTGC 6340  
 QY 1271 GCAAGCTGGAGGGCTGGAGCTGAGGACCCAGGACCTGGAGGCGCGCTGCGCACCAAGG 1330  
 Db 6341 AGCAGGAGCAGAGCGTCTGACACAGCTGCGCGGAGGCGGAGCGCGCGCGCGGCGG 6400  
 QY 1331 GCCTGGAGCTGGAGGTCTGTGAGAAATGAGTGCAGCGCAAGAAAGCAGCGGAGCTGC 1390  
 Db 6401 CTGAGGAGGCGGAGGAGCGCGCGGTGCGAGCGGAGCGTGAAGCGCGCAGCGCGCGCG 6460  
 QY 1391 TCGCGGAGAGAGTCAACCTCTGTCAGCAGGAGCTGTCAGGAGCTGCGGCGCGCGCGCG 1450  
 Db 6461 AGTGGAGAGGCGCGCGGCTGAAGCAGTGGCAGAGGAGCAGGCGCACAGCGCGCGGCTC 6520  
 QY 1451 TGGCGCGCGACATGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1510  
 Db 6521 AGGCACAGCGCGCTGCAGAGAGCTGCGCAAGGAGGCGCGCAGAGCGCGCGCGCGG 6580  
 QY 1511 TGGAGCGGCTGCGGCGCGAGCTGCGGAGGAGCGCGCAAGGCGCATGACAGATGCTCTCG 1570  
 Db 6581 CACAGCGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 6640  
 QY 1571 GCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGGAGAGAGTGTGTTGATACCAAGA 1630  
 Db 6641 ATAAGAAATTGCGCGAGC---AGACGCTGCGCGCAGAAAGGCGCGAGTGGAGCAGGAGCTGA 6697  
 QY 1631 AACAGCTGCAAGCAGAGCTACGTGGCCATGTACACGCGGAACCGCGCTGCGAGAGGCGCG 1690  
 Db 6698 CAACACTGCGGCTGCAGCTGGAGGAGAGCCGACACCAAGAGAGACTGCTGTGACGAGGAG 6757  
 QY 1691 TGCAGCAGCTGGCACGTGGGAGCAGCGCGGAGCGCTTGGAGGTTGACCTGGAAGGGG 1750  
 Db 6758 TGCAGCGGCTGAAGCGGAGGCGCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGAGG 6817  
 QY 1751 CTGACATCCC 1760  
 Db 6818 AGCTCTTCTC 6827

## RESULT 15

US-09-764-891-6987/c  
; Sequence 6987, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6987  
; LENGTH: 22585  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (9701)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-891-6987

Query Match 5.9%; Score 105.2; DB 9; Length 22585;  
Best Local Similarity 47.7%; Pred No. 4,4e-15;  
Matches 401; Conservative 0; Mismatches 434; Indels 6; Gaps 3;  
  
QY 915 GCTGAGGGCCGGAGCCCAAGCGGCAACAAGCTCAAGCAGGCTTCGACAGAGAGCCA 974  
Db 10432 GCAGGAGGGGACTGGGCTGCAGGAGGAGCGGGGCTGCAGGAGAGCGGGCTGCAGGA 10373  
  
QY 975 GCGCGCGCAGCAGGTCTCTGCACCTGCAGGTACTGACAGTTTCAGCAGGAGAGCGGCAGCT 1034  
Db 10372 GGGGACGGCGCTGCAGAGGAGATGCATGCAGAGGGGACGCGCAGATGCAGGAGGGGAC 10313  
  
QY 1035 CCGGCAGGAGCTCGAGAGCTCATGAAGCAGCAGGACCTGTGGAGACCAAGCTCAGGTC 1094  
Db 10312 CGAGTGCAGGAGGGGAGCGCAGATGCAGAGGGGACGGGGCTGCAGGAGGAGATGCAGGA 10253  
  
QY 1095 CTACGAGAGGGAAGAGACAGCTTCGGCCCGCGCTGAGGAGAGACCCAGTGGGAGGTGTG 1154  
Db 10252 GGGGCTGCAGGTACAGGAGGGGTTGGGGC--TGCAGGAGGGGACACAGATGCAGGAGGG 10196  
  
QY 1155 CCAGAGTTCAGGCGAGATCTCCTCTCTGAGCAGCAGCTGAGGAGTCCACAGCGAGGT 1214  
Db 10195 GCGCAGATGCAGAGGGGCGAGGGGCTGCAGAGGAGATGCAGATGCAGGAGGGGATGCA 10136  
  
QY 1215 GAACGCCAAGGTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGGGGCAA 1274  
Db 10135 GATGCAGGAGGAGATGCAGATGCAGAGGGGATGCAGGTGCAGGAGGGGACCGGGCTGCA 10076  
  
QY 1275 GCTGAGGGCTTGAGCTGAGACCCAGACCTTGAGGGCGCCCTGCGCAACCAAGGGCCT 1334  
Db 10075 GGAGGGAGCGGGCTTGGAGGGGACCGGGCTACAGGAGGAGATGCAGATGCAGGAGGGGT 10016  
  
QY 1335 GGAGCTGGAGGTCTGTGAGATGAGCTGCAGCGCAAGAGACAGCGGCGGAGCTGCTGG 1394  
Db 10015 GCACGTACAGGACTGGTTCGGGCTGCAGAGGGGACCGCAGATGCAGGAGGGCTCGGGCT 9956  
  
QY 1395 GGAGAGGTGAACCTGTGAGCAGGAGCTGCAGGAGCTGCGGGGCCAGAGCCCGCCCTGGC 1454  
Db 9955 GCAGGAGGGGAGGAGATCCAGGAGGGGACCGGGTGCAGGAGGGGACCGGGTGCAGGA 9896  
  
QY 1455 CCGCGACATGGGGCCGCCACCTTCCCGAGAGACCTCTCCCTGCGCTGCAGCGGGAGCTGGA 1514  
Db 9895 GGGGGCGCAGGGGCGAGGAGGAGCGGGTGCAGGAGGGGACACAGGTGCAGGAGGGGAC 9836  
  
QY 1515 GCGGCTGCGGGCCGAGCTGCGGAGGAGCGGCAAGCCATGACAGATGTCTCTGGGGCTT 1574  
Db 9835 GGGGCTGCAGGAGGAGATGCAGATCAGGGGGGTG-CAGGTACAGGAGGGGTTCGGGCT 9777  
  
QY 1575 CCAGCATGAGCGGCTCGTGTGAGAGGAGGAGAGAGGTGATTTCAGTACCAAGACA 1634

Db 9776 GCAGGAGGGGACGCAGATGCAGGAGGGGACGCAGATGCAGGAGGGGACACAGATGCAGAA 9717  
QY 1635 GCTGCAGCAGAGCTACTGTCGCCATGTATACAGCGGAACAGCGGCTTGAGAGAGGCCCTGCA 1694  
Db 9716 GGGGGCGCAGATGCANGAAGCGACGCGGGCTGCAGAGGAGATGCAGATGCAGGAAGGAT 9657  
QY 1695 GCAGCTGCACGTGGG--GACAGCGCGGGGAGCCCTTTGGAGGTTGACCTGGAAGGGCT 1752  
Db 9656 GCAGATCCCTGGAGGGATGCAGGTGCAGGAGGGGACGGGGCTGCAGGACGGGATGGGGCT 9597  
QY 1753 G 1753  
Db 9596 G 9596

Search completed: June 15, 2003, 06:49:17  
Job time : 263.718 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 14, 2003, 23:14:56 ; Search time 2493.59 Seconds  
(without alignments)  
11632.279 Million cell updates/sec

Title: US-09-513-888C-3  
Perfect score: 1791  
Sequence: 1 atgggcagcgcagtagcct.....tcatagccactgagatctga 1791

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_othr:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

AV663618 AV663618  
AW028197 wv84a01.x  
AW016544 UI-H-BIOP  
BF549120 UI-R-AO-a  
BF554804 UI-R-EO-C  
BE410921 601303579  
BE384131 601272956  
BF706253 280459 MA  
AZ858490 2M0163110  
AJ454524 AJ454524  
AV663617 AV663617  
BG750395 602703229  
BG079505 AGENCOURT  
BF037137 601456991  
BQ652556 AGENCOURT  
AL635635 AL635635  
BQ720762 AGENCOURT  
BQ652760 AGENCOURT  
BF970582 602274088  
BE740090 601595016  
BM947278 UI-M-EHOP  
BF312957 601896225  
BI854892 603381873  
AZ595691 1M0408A04  
BF784037 602107882  
BG117774 602818743  
BI520208 603071456  
BI338021 361664 MA  
BG255093 602369606  
BQ051027 AGENCOURT  
BE304332 601105731  
BE312971 601150210  
BQ955592 AGENCOURT  
BI752956 603025713  
BE757942 212520 MA  
AL551801 AL551801  
BF583930 602096941  
BE262105 601146482  
BE276168 601144424

ALIGNMENTS

RESULT 1  
BF058214/c 668 bp mRNA linear EST 16-OCT-2000

LOCUS 7k28d05.x1 NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:3476769 3'  
DEFINITION similar to TR:Q9Y5W1 Q9Y5W1 FEZ1. [5] TR:Q9Y5W0 TR:Q9Y5V9 TR:Q9Y5V8  
TR:Q9Y5V7 ;contains MER22.t3 TAR1 repetitive element ;, mRNA  
sequence.

ACCESSION BF058214  
VERSION BF058214.1 GI:10812110  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 668)  
REFERENCE NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento  
Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The  
I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	608.6	34.0	668	12 BF058214	BF058214 7k28d05.x
C 2	606.2	33.8	817	9 AI984777	AI984777 wr85b07.x
C 3	545	30.4	545	10 AW007737	AW007737 wr68e06.x
C 4	523.4	29.2	740	14 BQ769435	BQ769435 UI-M-F10-
C 5	489.6	27.3	547	12 BF590813	BF590813 7h42e07.x
C 6	447	25.0	447	9 AI042490	AI042490 ox62e04.x

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Seq primer: -40UP from Gibco  
High quality sequence stop: 441.  
Location/Qualifiers  
1. 668

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3476769"  
/clone\_lib="NCI CGAP Ovi8"  
/tissue\_type="Fibrotheoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAGTGGGAGCGCGGACATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 105 a 185 c 226 g 150 t 2 others  
ORIGIN  
Query Match 34.0%; Score 608.6; DB 12; Length 668;  
Best Local Similarity 95.2%; Pred. No. 3.9e-101;  
Matches 637; Conservative 0; Mismatches 31; Indels 1; Gaps 1;  
QY 190 GACTCTTCTACATCAAGTGCAGCAAGAGCGCGGCTCCATCCAGATACAG 249  
DB 668 GACTCTTCTTAAATCAATGTAGCCAGTAAGCCGAGGCTCCATCCCAAGATTTATCG 609  
QY 250 GCATGTCCAGCGGGATTAGGGGGCCAGGCTGGGGTGGATTTGACCGGTCCACACCC 309  
DB 608 GCATGTCCAGCGGAATAGAGGGCCAGATGGGGTGAATTTGACNCGTCCACCA 549  
QY 310 CCCAGCTCATGCCCCCTTCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGGTGCAGTGAAG 369  
DB 548 CCCAGCTCATGCCCC-TCTCCAAATCAGCTAGAAATGGGCTCCGAGAGGGTGCAGTGAAG 490  
QY 370 CCCAGAGCTTCAAGCCTGTGTGCACAGGTGCAGAGGCATCTGTCATCTCTCCCGGAG 429  
DB 489 CCCAGAGCTTCAAGCCTGTGTGCACAGGTGCAGAGGCATCTGTCATCTCTCCCGGAG 430  
QY 430 AGTGCCAGCCACAGCTGCACCCGCCCTCCAGACAAGCCCAAGGACGAGCTGAAG 489  
DB 429 AGTGCCAGCCACAGCTGCACCCGCCCTCCAGACAAGCCCAAGGACGAGCTGAAG 370  
QY 490 CTGCGCCTGTGCTCTGGGGCGCTGTGCAGCTCCCGCGGAACTCCATGTCAGGCTGCC 549  
DB 369 CTGCGCCTGTGCTCTGGGGCGCTGTGCAGCTCCCGCGGAACTCCATGTCAGGCTGCC 310  
QY 550 ACACACAGCAGCAGCAGCTACAGCTGGACCGCTGGTTCACACCGTGGACCCACA 609  
DB 309 ACACACAGCAGCAGCAGCTACAGCTGGACCGCTGGTTCACACCGTGGACCCACA 250  
QY 610 AGCCGCTTTGGGGGCTCCGCCCAACAATCACCCAGGCGATCGTCTCCAGCAGCAGCAAC 669  
DB 249 AGCCGCTTTGGGGGCTCCGCCCAACAATCACCCAGGCGATCGTCTCCAGCAGCAGCAAC 190  
QY 670 ATGATGAGCCTGAAGGCTCTGTGCTTCTCCGAGGAGGTAGCAAGCTGGGCGCACTCGAAC 729  
DB 189 ATGATGAGCCTGAAGGCTCTGTGCTTCTCCGAGGAGGTAGCAAGCTGGGCGCACTCGAAC 130  
QY 730 AAGGAGACAGAGGGCCCTCTGTGTCTCGCTCCCGCTCCCGCTCCAGCAGCAGTGCAGCATC 789  
DB 129 AAGGAGACAGAGGGCCCTCTGTGTCTCGCTCCCGCTCCCGCTCCAGCAGTGCAGCATC 70  
QY 790 CAGGAGCTGGAGCAGAACTGTGTGAGAGGAGGAGCGCCCTCCAGAACTGCAGGCGCAG 849  
DB 69 CAGGAGCTGGAGCAGAACTGTGTGAGAGGAGGAGGCGCCCTCCAGAACTGCAGGCGCAG 10  
QY 850 TTTGAGGAG 858  
DB 9 TTTGAGGAG 1

RESULT 2  
AI984777/c  
LOCUS  
DEFINITION  
AI984777.1 GI:5812054  
ACCSSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 817)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 1307 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 428.  
Location/Qualifiers  
1. 817

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2494453"  
/clone\_lib="NCI CGAP\_Kid11"  
/lab\_hosts="DH10B"  
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 133 a 235 c 266 g 180 t 3 others  
ORIGIN  
Query Match 33.8%; Score 606.2; DB 9; Length 817;  
Best Local Similarity 94.0%; Pred. No. 1.1e-100;  
Matches 717; Conservative 0; Mismatches 35; Indels 11; Gaps 8;  
QY 112 TCCGACGGGCTGTGAGGTTGGCTTCTCCAGGACTCCGGTCACGCAAGTCCAGTCC 171  
DB 756 TCCGACGGGCTGTGAGGTT--GCTTTTCCCGGACTCC-GTCACGCAAGTCCAGTCC 700  
QY 172 AAAATGGGCAAGCGCAAGACTTCTTACATCAAGGTTCAGCCAGAAAGCCCGGGCTCC 231  
DB 699 AAA--TGGCAGAGCGAGACTTCTTACATCAAGG-CAGCCAGAAAGCCCGGGCTCC 643  
QY 232 CATCACCAGATTACCGGCACTGTCCAGCGGGGATTAGGGGCCAGGCT-GGGGTGA 290  
DB 642 CATCACCAGATAACAC-GCACGGTCCAGCGGGATTAGGGGCCAGGTTGGGGGTGA 584  
QY 291 CTTTGACCCCTCCACACCCCGGAGCTCATGCCCTTCTCCAAATCAGTAGAAATGGGCTC 350  
DB 583 CTNTGCCCGGCTCCACCCCGGAGGCTCATGCCCTTCTCCAAATCAGTAGAAATGGGCTC 524  
QY 351 CGAAGAGGCTGAGTGAGGCCACACAGGCTTTCAGGCTGTGCTGCCAGGTCAGGAGCCAT 410

523	Db	GTGAAGGGTGAGTAAGGCCCCAGAGCCTTCAAGCTTGTGTCACGGTTCAGAGCCAT	464
411	Qy	CCTGCACCTCTCCCGGAGAGTGCAGCCACCACTGCACCCCGCCCTCCAGACAAGCC	470
463	Db	CCTGCACCTCTCCCGGAGAGTGCAGCCACCACTGCACCCCGCCCTCCAGACAAGCC	404
471	Qy	CAAGGAGCAGGAGCTGAAGCTGGCGTGTGCTCTGGGGCGCTGTCAAGCTTCGGCCGAA	530
403	Db	CAAGGAGCAGAAGCTGAAGCCTGGCGTGTGCTCTGGGGCGCTGTCAAGCTTCGGCCGAA	344
531	Qy	CTCCATGTCCAGCGTGCACCAACAG - CACCAGCAGCTACCACT - -GGACCCGCT	587
343	Db	CTCCATGTCCAGCGTGCACCAACAGCCCGTGTACCAAGTTGGGACCCGCT	284
588	Qy	GGTCACACCCGTTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCAACAACATCACCCAGG	647
283	Db	GGTCACACCCGTTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCAACAACATCACCCAGG	224
648	Qy	CATCGTCTCCAGGACAGCAACATGATAGCGCTGAAGGCTCTGTCTTCTCCGACGGAG	707
223	Db	CATCGTCTCCAGGACAGCAACATGATAGCGCTGAAGGCTCTGTCTTCTCCGACGGAG	164
708	Qy	TAGCAAGCTGGSCACTCGAAACAAGGACAGCAAGGGCCCTCTGTGTGTCGCTCCCCCAT	767
163	Db	TAGCAAGCTGGSCACTCGAAACAAGGACAGCAAGGGCCCTCTGTGTGTCGCTCCCCCAT	104
768	Qy	CTCCACGAGTGTCAGCATCCAGGAGCTGGAGCAGAAGCTGTGGAGGGAGGGCGC	827
103	Db	CTCCACGAGTGTCAGCATCCAGGAGCTGGAGCAGAAGCTGTGGAGGGAGGGCGC	44
828	Qy	CCTCAGAAGCTGCAGCGCAGCTTTGAGGAGGAAGGAGCTTGCC	870
43	Db	CCTCAGAAGCTGCAGCGCAGCTTTGAGGAGGAAGGAGCTTGCC	1

RESULT 3	
AW007737/c	
LOCUS	
DEFINITION	
<hr/>	
	AW007737          545 bp     mRNA        linear    EST 09-MAR-2000
	wc68606.xl Soares thymus_NHPTH Homo sapiens cDNA clone
	IMAGE:2512642 3' similar to contains MER22.t3 TARI repetitive
	element ; mRNA sequence.

ACCESSION AW007737  
 VERSION AW007737.1  
 KEYWORDS EST.  
 SOURCE human.

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 545)  
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strauch, Ph.D.

COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1333 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 463.

## FEATURES

```

/clone_lib="Soares_thymus_NHFTb"
/dev stage="fetal"

```

```

/dev_scages="lctal"
/lab_host="DH10B (phage-resistant)"
/note="Organ: thymus, pooled; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5,

```

TCTTACCAATCTGAATGGAGCGCGCCGCAAGCTTTTTTTTTTTTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p773 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bonaldo. "  
 80 a 159 c 189 q 117 t

BASE COUNT	80 a	159 c	189 g	117 t	
ORIGIN					
Query Match	30.4%	Score 545	DB 10	Length 545	
Best Local Similarity	100.0%	Pred. No. 1.5e-89			
Matches 545; Conservative	0	Mismatches 0	Indels 0	Gaps 0	
QY	326	TTCTCCAAATCAGCTAGAAATGGGCTTCGAGAAAGGGTGCACTGAGGCCACAGCCTTCAAGC	385		
Db	545	TCTCCAAATCAGCTAGAAATGGGCTTCGAGAAAGGGTGCACTGAGGCCACAGCCTTCAAGC	486		
QY	386	CTGTGCTGCGACGGTCAGGAGGCATCTCTCACTCTCTCCCGGAGAGTGCCAGGCCACGAC	445		
Db	485	CTGTGCTGCGACGGTCAGGAGGCATCTCTCACTCTCTCCCGGAGAGTGCCAGGCCACGAC	426		
QY	446	TGCACCCCGCCCTTCAGACAAGCCCAAGGACGAGAGCTGAAGCCTGSCCTGTGTCTCTG	505		
Db	425	TGCACCCCGCCCTTCAGACAAGCCCAAGGACGAGAGCTGAAGCCTGSCCTGTGTCTCTG	366		
QY	506	GGGCGGTGTTCAGACTCCGGCCGGAACTCCATGTCNAGCTGCCCCACACACAGCACACGCA	565		
Db	365	GGGCGGTGTTCAGACTCCGGCCGGAACTCCATGTCNAGCTGCCCCACACACAGCACACGCA	306		
QY	566	GCAGCTACCACTGGACCCCGCTGGTTCACACCGTGGGACCCACAGCCGCTTTTGGGGGCT	625		
Db	305	GCAGCTACCACTGGACCCCGCTGGTTCACACCGTGGGACCCACAGCCGCTTTTGGGGGCT	246		
QY	626	CGCGCCACAAATCAACCAGGGCATCGTCTCTCAGGACAGCAACATGATGAGCCTCAAGG	685		
Db	245	CGCGCCACAAATCAACCAGGGCATCGTCTCTCAGGACAGCAACATGATGAGCCTCAAGG	186		
QY	686	CTCTGTCTTTTCGACGGAGGTAGCAAGCTGGGCCACTTCGAACAAGGACAGCAAGGGCC	745		
Db	185	CTCTGTCTTTTCGACGGAGGTAGCAAGCTGGGGCCACTTCGAACAAGGACAGCAAGGGCC	126		
QY	746	CCTCGTGTGTCGCTCCCCCATCTCCACGGACGAGTGACGATCCAGGAGCTGGAGCAGA	805		
Db	125	CCTCGTGTGTCGCTCCCCCATCTCTCCACGGACGAGTGACGATCCAGGAGCTGGAGCAGA	66		
QY	806	AGCTGTTGGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTTGAGGAGAAAGCAGC	865		
Db	65	AGCTGTTGGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTTGAGGAGAAAGCAGC	6		
QY	866	TTGGCC	870		
Db	5	TTGCC	1		

RESULT 4	BQ769435	740 bp	mRNA	linear	EST 26-JUL-2002
LOCUS	BQ769435				
DEFINITION	UI-M-F10-byq-a-22-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone IMAGE:6400797 5', mRNA sequence.				

ACCESSION BQ769435  
 VERSION BQ769435.1 GI:21977909  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;  
1 (bases 1 to 740)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

## FEATURES

Seq primer: pYX-5,  
Location/Qualifiers  
1..740

/organism="Mus musculus"

/strain="CS7BL/6"

/db\_xref="taxon:10090"

/clone\_lib="NIH BMAP F10"

/tissue\_type="whole brain"

/dev\_stage="embryo 12.5dpc"

/lab\_host="DH10B (T1 phage resistant)"

/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CAGCCAGGAC. This library was created for the University  
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

## BASE COUNT

185 a 213 c 222 g 120 t

## ORIGIN

Query Match 29.2%; Score 523.4; DB 14; Length 740;

Best Local Similarity 84.1%; Pred. No. 1.3e-85;

Matches 628; Conservative 0; Mismatches 111; Indels 8; Gaps 3;

QY 425 CGGAGAGTCCAGCCAGCAGCTGACCCCGCCCTCCAGACAGCCAGGAGCAGGAGC 484  
DB 2 CTGAGAGTACAGCCAGCAGCAACTTCATCCCTCCAGATAAGCCAGGAGCAGGAGC 61  
QY 485 TGAAGCCTGGCTGTGCTCTGGGGCGCTGTGAGACTCCGCGGAACTCCATGTCAGCC 544  
DB 62 TGAAGCCAGGCTGTGCTCGGGGCACTGTCTGACTCAGCGCGAACTCCATGCTAGCC 121  
QY 545 TGCCACACACAGCAGCAGCAGCTACCGAGTGGACCCGCTGTGTCACCCGCGGAC 604  
DB 122 TGCCACAGCAGCAGCAGCAGCAGCTACCGAGTGGATCTCTGTGTCACCCCGGGGC 181  
QY 605 CCACAGCGCTTTTGGGGCTCGGCCCAACATCACCCAGGCGATCGCTCCAGGACA 664  
DB 182 CTACAGCGCTTTGGGGTTCAGCTCACAAATCACAGGCAATCATCTTCAGGACA 241  
QY 665 GCAACATGATGAGCTGAGGCTGTGCTTCTTCGAGGAGGTAGAGCTGGGCCACT 724  
DB 242 GTAATATGATGAGCTGAGGCTGTGCTTCTGTGATGGCGCAGCAAGCTGGCTACC 301  
QY 725 CGAACAGGCGACAGAGGCGCCCTCGTGTGTCGCTCCCATCTCCAGGACGAGTGCA 784  
DB 302 CAGGCAAGCAGATAGGCGGCTCTGTGTGCTGTCCCACTCTCCAGGATGATGCA 361  
QY 785 GCATCCAGGAGCTGGAGCAGAGAGCTGTTGGAGAGGAGGCGGCCCTCCAGAGCTGCAGC 844  
DB 362 CCATCCAGGAGCTGGAGCAGAGAGCTGCTGCAGAGGAGAGCTGCATCAGAGAGCTACAGC 421  
QY 845 GCAGCTTTGAGGAGAGGAGCTTGCTTCAGCTCCAGCTGGCCTACAGAGGCGCGCGGCT 904  
DB 422 GCAGCTTTGATGAGAGGAGTTTGCTCTCTGGCCAAACCTTTGAGAGGCGGCCACGGCGGA 481

QY 905 GCAGGAGCAGCTGGAGGGCCCGAGAGCCCAAGCGGCAACAGCTCAAGCAGGCTCGC 964  
DB 482 CTAGAGACGACTGGAGTGGCTGGAGCTAA-----GAGCAAGCTGAGCCACCTCAC 535  
QY 965 AGAAGACCGAGCGCGCCAGCAGAGCTCTGCACCTGCGAGGTACTGACGCTTCAGCAGGAGA 1024  
DB 536 AGAAGACCGAGCGCGCACACAGCAGGTGCTCCAGCTGCGAGGTGTTGTCAGTTGCGAGGAGA 595  
QY 1025 AGCGGAGCTCCGCGAGGAGCTCGAGAGCTCTGAGAGGAGCAGGAGCTGCTGAGAGCA 1084  
DB 596 AACGGCAGCTAAGCGAGGAGCTAGAGAGCTTAATGAAGGAGCAGGAGCTGCTGAGAGCA 655  
QY 1085 AGCTCAGGTCTTACGAGAGGAGAGCAGCTTCGCGCCCGCTGAGAGGAGACCCAGT 1144  
DB 656 AGCTGAGGTCTTACGAGCGGGAGAA-ACCACTTCGCGCCCTGCACTGGAGGAGA-CCAT 713  
QY 1145 GGGAGGTGTGCCAGAACTCAGGCGAGA 1171  
DB 714 GGGAGGTGTGCCAGAGTCAGTGCAGA 740

RESULT 5  
BF590813/c  
LOCUS  
DEFINITION 7h42e07.x1 NCI CGAP Col6 Homo sapiens cDNA clone IMAGE:3318660 3,  
similar to TR:Q9Y5W1, Q9Y5W1, FEZ1, [5] TR:Q9Y5W0, TR:Q9Y5V9, TR:Q9Y5V8,  
TR:Q9Y5V7; contains MER22.t3 TARI repetitive element ;, mRNA  
sequence.  
ACCESSION BF590813  
VERSION BF590813.1 GI:11683137  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 547)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck, M.D.,  
Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 434.  
Location/Qualifiers  
1..547  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3318660"  
/clone\_lib="NCI CGAP Col6"  
/tissue\_type="colon tumor, RER+"  
/lab\_host="DH10B"  
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; Site 1: Not I; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI CGAP Col6 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(clones 1057416-1061255, and 1144584-1145351).  
Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 83 a 156 c 189 g 119 t

ORIGIN

Query Match 27.3%; Score 489.6; DB 12; Length 547;

/notes="Organ: mixed (see below); Vector: )pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NBHPu, and fetal heart NBHH19w) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

Db	147	TCGAACAAGGCAGACAAGGGCCCCCTGGTGTGTCGCTCCCCCATCTCCACGGACGAGTGC	88
Qy	784	AGCATCCAGGAGCTGGAGCAGAGAAGCTGTTGGAGAGAGGGGGCCCCCTCCAGAAGCTGCAG	843
Db	87	AGCATCCAGGAGCTGGAGCAGAGAAGCTGTTGGAGAGAGGGGGCCCCCTCCAGAAGCTGCAG	28
Qy	844	CCGAGCTTTGAGGAGAAGAGGCTTCCC	870
Db	27	CCGAGCTTTGAGGAGAAGAGGCTTCCC	1
RESULT	7		
LOCUS	AV663618		
DEFINITION	AV663618 Bos taurus brain fetus Bos taurus cDNA clone E1BR037F07	596 bp	linear
ACCESSION	AV663618		
VERSION	AV663618.1		
KEYWORDS	EST.		
SOURCE	COW.		
ORGANISM	Bos taurus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.		
AUTHORS	1 (bases 1 to 596)		
TITLE	Takasuka,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H. and Sugimoto,Y.		
	Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000		
JOURNAL	Bovine ESTs		
MEDLINE	Nucleic Acids Res. 29 (22), E108 (2001)		
COMMENT	21570554		
	Contact: Yoshikazu Sugimoto		

Animal Genetics Division  
Shirakawa Institute of Animal Genetics  
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan  
Tel: 81-248-25-5641  
Fax: 81-248-25-5725  
Email: kazusugi@ccococ.ocn.ne.jp  
Single pass sequencing.  
This clone was obtained from a polyA-deleted cDNA library.

# FEATURES

Location/Qualifiers  
1..596  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone="E1BR037F07"  
/clone\_lib="Bos taurus brain fetus"  
/tissue\_type="brain"  
/dev\_stages="fetus"  
/lab\_host="DH10B"  
/note="Vector: pZU1; Site 1: SalI; Site 2: NotI; Poly A was deleted from a NotI site"  
114 a 196 c 170 g 103 t 13 others

Query Match 23.0%; Score 412.2; DB 10; Length 596;  
Best Local Similarity 87.9%; Pred. No. 2.3e-65;  
Matches 471; Conservative 0; Mismatches 61; Indels 4; Gaps 2;  
QY 92 TCACAAAGCTCAACGGTATTCCGACGGCTGCTGAGTTGGCTTCTCCAGGACTCG 151  
DB 61 TCACAAAGCTCAACGGTATTCCGACGGCTGCTGAGTTGGCTTCTCCAGGACTCG 120  
QY 152 GTCACGGAAGTCACAGTCCAAATATGGCAAGAGCGAAGACTTCTTACATCAAGGTCA 211  
DB 121 GACATGGCAAGTCCAGTCCAAATATGGCAAGAGCGAAGACTTCTTACATCAAGGTCA 180  
QY 212 GCCAAGAACCCGGGCTCCCATCATCACCGCACTATACACGGCACTGCCAGCGGGATTAG 271  
DB 181 GCCAAGAGCCCGGGCTCCCATCATCACCGCACTATACACGGCACTGCCAGTGGGNNNTAG 240  
QY 272 GGGGCCAGCTGGGTGACCTTTGACCGGTTCACACCCCGCCAGCTCATGCCCTTCTCA 331  
DB 241 GGGGCCAGCTGGGTGACCTTTGACCGGTTCACACCCCGCCAGCTCATGCCCTTCTCA 300  
QY 332 ATCAGCTAGAAATGGCTCCGAGAAGGTGACGTGAGGCGCCACAGCTTCAAGCTGTGC 391  
DB 301 ATCAGCTAGAGATGGTCTGAGAGAGGTGCTGAGACCCACAGCTTCAAGCGGTGC 360  
QY 392 TGCCACGCTCAGGAGCCATCTTGCATCTCTCCCGAGAGTGCCAGCCACAGCTGCACC 451  
DB 361 TGCCACGCTCAGGAGCCATCTTGCATCTCTCCCGAGAGTGCCAGCCACAGCTGCATC 420  
QY 452 CCGCCCTCCACAGACCCAGGAGGAGGAGTCAAGCTGAGCTGCTGTCTGGGGCGC 511  
DB 421 CTGCGCTCTTGACAAAGCCAGGAGGAGGAGTCAAGCTGAGCTGCTGTCTGGGGCGC 480  
QY 512 TGTCAGACTCCCGCGGAACTCCATGTCAGCTGCGCCACACACAGCACCAGCAGCT 571  
DB 481 TGTCAGACTCCCGCGGAACTCCATGTCAGCTGCGCCACACACAGCACCAGCAGCT 540  
QY 572 A-CCAGCTGACCCGCTGG---TCACACCCGTGGGACCCACAAAGCCGTTTGGGGG 623  
DB 541 ACCCAGCTGACCCGCTGGTGTCACTCTGTGGGGGCGCCCGCCAGCCGCTTGGGGG 596

RESULT 8  
AW028197/c  
LOCUS  
DEFINITION  
IMAGE:2536200 3' similar to TR:060299 O60299 KIAA0552 PROTEIN. ;  
mRNA sequence.  
AW028197  
AW028197.1 GI:5886953  
EST.  
SOURCE  
human.

# ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

# REFERENCE

1 (Bases 1 to 397)

# AUTHORS

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

# TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

# JOURNAL

Tumor Gene Index

# COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Insert Length: 1067 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 139.

# FEATURES

Location/Qualifiers

1..397

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2536200"

/clone\_lib="Soares\_thymus\_NHFTb"

/dev\_stage="fetal"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: thymus, pooled; Vector: pT73D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5,

TGTACCAACTGAAGTGGAGCGCGCGACAGTTTCTTTTCTTTT 3'],

double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo.

67 a 120 c 124 g 86 t

# BASE COUNT

# ORIGIN

Query Match 20.5%; Score 366.6; DB 10; Length 397;  
Best Local Similarity 95.2%; Pred. No. 4.3e-57;  
Matches 378; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
QY 502 TCTGGGGCGCTGTGACACTCCGGCCGGAATCTCCATGTCCAGCCTGCCACACAGCACC 561  
DB 397 TCTGGGGAGCTGTGAGAATCCGCTGGAAATTCATGTCCAGCCTGCCACACAGCACA 338  
QY 562 AGCAGCAGTCACTAGCTGGAGCCCGCTGTCACACCCGTGGGACCCACAGCCGTTTGGG 621  
DB 337 GGCAGCAGTCACTAGTGGACCCGCTGTCACACAGTGGGACCCACAGCCGTTTGGG 278  
QY 622 GGCTCCGCCCAACACATCACCAGGGCATCTCTCTCCAGGACAGCAACATGATGAGCCTG 681  
DB 277 GGCTCCGCCCAACACATCCTCAGGGCATCTCTCTCCAGGACAGCAACATGATGAGCTT 218  
QY 682 AAGGCTCTGCTCTCTCTCTCGAGCGAGGTAGCAAGCTGGGCCACTCGAACAGGACAGCAAG 741  
DB 217 AAGGCTCTGCTCTCTCTCTCGAGCGAGGTAGCAAGTTGGGCCACTCGAACAGGACAGCAAG 158  
QY 742 GGCCCTCTGCTGTCGCTCCCTCCATCTCCAGGACGAGTCCAGCATCCAGGAGTGGAG 801  
DB 157 GGCCCTCTGCTGTCGCTCCCTCCATCTCCAGGATGAGTGGAGCATCCAGGAGTGGAG 98  
QY 802 CAGAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAGCTGTCAGCGCAGCTTTGAGGAGAG 861  
DB 97 CAGAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAGCTGTCAGCGCAGCTTTGAGGAGAG 38  
QY 862 GAGCTTGCTCCAGCTTGGCTTACGAGGAGCGGCGC 898  
DB 37 GAGCTTGCTCCAGCTTGGCTTACGAGGAGCGGCGC 1

# RESULT 9

AW016544/c

# LOCUS

DEFINITION  
UI-H-B10p-abg-g-04-0-UI.s1 NCI\_CGAP\_Sub2 Homo sapiens cDNA clone  
AW016544  
344 bp mRNA linear EST 10-SEP-1999



```

IMAGE:2711982 3', mRNA sequence.
AW016544
VERSION AW016544.1 GI:5865301
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 344)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: NCI-CGAP clone distribution
Information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=No.

FEATURES
    Location/Qualifiers
        1..344
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:2711982"
            /clone_lib="NCI CGAP Sub2"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pT73D-Pac (Pharmacia) with a modified
            polylinker; Site 1: Not 1; Site 2: Eco RI; The
            NCI CGAP Sub2 library is a subtracted library derived from
            B1. B1 constitutes a mixture of 21 normalized or
            subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP Pr22
            , NCI CGAP Pr28, NCI CGAP Co10, NCI CGAP Co16,
            NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3,
            NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8,
            NCI CGAP CLL1, NCI CGAP Lei2, NCI CGAP Brn23, NCI CGAP Lus
            , NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP GC4, NCI CGAP GC6
            , NCI CGAP Brn25. These 21 libraries were pooled and a
            single-stranded DNA preparation of the resulting mixture
            was used as a tracer in a subtractive hybridization with a
            driver whose composition is detailed below: NCI CGAP Kid3
            pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE
            CloneIDs 132376-132391, 1456008-1456775, 1500552-1502855
            ) NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725,
            3776-3778 (IMAGE CloneIDs 1323912-1325831,
            1473368-1472903, 1492104-1493255) NCI CGAP Lus pool 1 LLAM
            3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991,
            1520904-1522439) NCI CGAP GC4 pool 1 LLAM 3164-3167,
            3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631,
            1469064-1470983, 1475592-1476743) NCI CGAP Pr22 pool 1
            LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs
            985608-986759, 1101192-1101959, 1217928-1220615)
            NCI CGAP Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE
            CloneIDs 1057416-1061255, 1144584-1145351) The resulting
            subtracted library contained 4 million recombinants.
            Subtraction was performed as previously described [Bonaldo
            , Lennon & Soares (1996): Normalization and Subtraction:
            Two Approaches To Facilitate Gene Discovery. Genome
            Research 6, 791-806.
            TAG LIB=NCI CGAP GC4
            TAG TISSUE=Term Cell
            TAG_SEQ=AAATC"
        51 a 106 c 112 g 74 t 1 others
BASE COUNT 51 a 106 c 112 g 74 t 1 others
ORIGIN
Query Match 18.9%; Score 338.2; DB 10; Length 344;
Best Local Similarity 98.8%; Pred. No. 6,3e-52;
Matches 340; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
554 ACAGCACCAGCAGCAGCTTACCGCTGGACCGCTGGTCCACACCCGTGGGACCCCAAGCC 613
|||||

```

```

Db 344 ACAGCACCAGCAGCAGCTTACCGCTGGACCGCTGGTCCACACCCGTGGGACCCCAAGCC 285
Qy 614 GTTTTGGGGGCTCGCGCCACAAATCACCAGGGCATCGTCTCTCCAGGACAGCAACATGA 673
Db 284 GTTTTGGGGGCTCGCGCCACAAATCACCAGGGCATCGTCTCTCCAGGACAGCAACATGA 225
Qy 674 TGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCCACCTCGAACAGG 733
Db 224 TGAGCCTGAAGGCTCTGTCTTCTCCGACGAGGTAGCAAGCTGGGCCACCTCGAACAGG 165
Qy 734 CAGACAGGGGCCCTCGTGTCTCGCTCCGCCATCTCCACGACAGTGCAGCATCCAGG 793
Db 164 CAGACAGGGGCCCTCGTGTCTCGCTCCGCCATCTCCACGACAGTGCAGCATCCAGG 105
Qy 794 AGCTGGAGCAGAAGCTGTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGGCACGCTTTG 853
Db 104 AGCTGGAGCAGAAGCTGTGTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGGCACGCTTTG 45
Qy 854 AGGAGAGGAGCTTGCCTCCAGCTGGCTTACGAGGAGCGGCCG 897
Db 44 AGGAGAGGAGCTTGCCTCCAGCTGGCTTACGAGGAGCGGCCG 1

RESULT 10
BF549120
LOCUS
DEFINITION UI-R-A0-af-c-04-0-UI.r1 UI-R-A0 Rattus norvegicus cDNA clone
UI-R-A0-af-c-04-0-UI 5', mRNA sequence.
ACCESSION BF549120
VERSION BF549120.1 GI:11640185
KEYWORDS EST.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 394)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cDNA Library Preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1791781
Seq primer: M13 Forward.
Location/Qualifiers
    1..394
        /organism="Rattus norvegicus"
        /strain="Sprague-Dawley"
        /db_xref="taxon:10116"
        /clone="UI-R-A0-af-c-04-0-UI"
        /clone_lib="UI-R-A0"
        /dev_stage="adult"
        /lab_host="DH10B (Life Technologies)"
        /note="Vector: pT73D-Pac (Pharmacia) with a modified
        polylinker; Site 1: Not 1; Site 2: Eco RI; This library
        consists of a mixture of individually tagged normalized
        libraries constructed from rat placenta, adult lung, brain
        , liver, kidney, heart, spleen, ovary, and muscle. The tag
        is a string of 3-5 nucleotides present between the Not I
        site and the oligo-dr track which allows identification of
        the library of origin of a clone within the mixture."
BASE COUNT 89 a 125 c 103 g 77 t
ORIGIN

```



```

source
1. .743
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3638203"
/clone_lib="NIH_MGC_21"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      119 a      229 c      226 g      168 t      1 others
ORIGIN
Query Match      16.7%; Score 298.4; DB 10; Length 743;
Best Local Similarity 97.3%; Pred. No. 1.3e-44;
Matches 324; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Qy      622 GGCTCCGCCCAACATCACCAGGCGATCGTCTCCAGGACAGCAACATGATGAGCCTG 681
      |||
Db      711 GGGCCGCCGCCAGCAATCACAGG-CATGCTCTCCAGGACGCAACATGATGAGCCTG 653

Qy      682 AAGGCTGTGCTTCTCCGAGGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACACAAG 741
      |||
Db      652 AA-GCTGTGCTTCTCCGAGGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACACAAG 594

Qy      742 GGGCCCTGTGTGCTGCTCCCTCCCATCTCCAGGACGAGTGCGATCCAGGAGCTGGAG 801
      |||
Db      593 GGGCCCTGTGTGCTGCTCCCTCCCATCTCCAGGACGAGTGCGATCCAGGAGCTGGAG 534

Qy      802 CAGAAGCTGTTGGAGAGGAGGGCCCTCCAGAGAGTGCAGCGAGCTTTGAGGAGAAG 861
      |||
Db      533 CAGAAGCTGTTGGAGAGGAGGGCCCTCCAGAGAGTGCAGCGAGCTTTGAGGAGAAG 474

Qy      862 GAGCTTGCTCCAGCTGCGCTTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAG 921
      |||
Db      473 GAGCTTGCTCCAGCTGCGCTTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAG 414

Qy      922 GGGCCGAGGCCAAAGGGCGCAACAGCTCAAG 954
      |||
Db      413 GGGCCGAGGCCAAAGGGCGCAACAGCTCAAG 381

RESULT 13
BE384131/c
LOCUS      BE384131      673 bp      mRNA      linear      EST 21-JUL-2000
DEFINITION 601272956F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614076 5',
mRNA sequence.
ACCESSION  BE384131
VERSION     BE384131.1  GI:9329496
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 673)
AUTHORS   NIH-MGC http://mgc.ncl.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTF
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCN275 row: P column: 13
            High quality sequence stop: 670.

FEATURES
source
1. .673
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3614076"
/clone_lib="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      110 a      211 c      208 g      144 t
ORIGIN
Query Match      16.0%; Score 286.8; DB 10; Length 673;
Best Local Similarity 98.7%; Pred. No. 1.7e-42;
Matches 310; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy      642 CCAGGGCATCGTCTCCAGGACAGCAACATGATGAGCTCTGTCTCTCTCCGA 701
      |||
Db      669 CCCAGGCATCGTCTCTCA-GACAGCAACATGATGAGCTCTGTCTCTCTCCGA 611

Qy      702 CGGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACAGAGGG-CCCTCTGTGTGCTCGCT 760
      |||
Db      610 CGGAGGTAGCAAGCTGGGCCACTCGAACAGGCGACAGAGGGCCCTCTGTGTGCTCGCT 551

Qy      761 CCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTTTGGAGAGGG 820
      |||
Db      550 CCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAGCTTTGGAGAGGG 491

Qy      821 AGGGGGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGG 880
      |||
Db      490 AGGGGGCCCTCCAGAGCTGCAGCGCAGCTTTGAGGAGAGAGCTTCCCTCCAGCTGG 431

Qy      881 CCTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGAGGCCCAAAGGCG 940
      |||
Db      430 CCTACGAGGAGCGCGCGCGCTGCGAGGACGAGCTGGAGGCGCGAGGCCCAAAGGCG 371

Qy      941 GCAACAAGCTCAAG 954
      |||
Db      370 GCAACAAGCTCAAG 357

RESULT 14
BF706253
LOCUS      BF706253      472 bp      mRNA      linear      EST 25-APR-2001
DEFINITION 280459 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION  BF706253
VERSION     BF706253.1  GI:11997851
KEYWORDS   EST.
SOURCE     cow.
ORGANISM   Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 472)
AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrnenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keefe,J.W.
TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
JOURNAL   Genome Res. 11 (4), 626-630 (2001)
MEDLINE   21180013
COMMENT   Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA

```

Tel: 402 762 4366  
Fax: 402 762 4390

Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCAGGAG

Plate: 77 row: A column: 4

Seq primer: ATTAGGTGACACTATAG.

#### FEATURES

source

1. 472

/organism="Bos taurus"

/db\_xref="taxon:9913"

/clone\_lib="MARC 3BOV"

/tissue\_type="pooled"

/lab\_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal

longissimus muscle."

92.a 171 c 123 g 86 t

Query Match 16.0%; Score 286; DB 12; Length 472;

Best Local Similarity 92.3%; Pred. No. 2.2e-42;

Matches 301; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ATGGCAGCGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60

Db 147 ATGGCAGCGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 206

Qy 61 TCGCAGTACAAGTCGCGAAGTCCTCCACCTCAAGAAAGCTCAACCGGTATTCGACGG 120

Db 207 TCGCAGTACAAGTCGCGAAGTCCTCCACCTCAAGAAAGCTCAACCGGTATTCGACGG 266

Qy 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCGGTCAGCGCAAGTCCAGCTCCAAATGGC 180

Db 267 CTGCTGAGTTGGCTTCTCCAGAGCTCCGGTCAGCGCAAGTCCAGCTCCAAATGGC 326

Qy 181 AAGCGCAAGACTTCTTACATCAAGTTCAGCCAGAAAGCCGGGCTCCATCACCCA 240

Db 327 AAGCGCAAGACTTCTTACATCAAGTTCAGCCAGAAAGCCGGGCTCCATCACCCA 386

Qy 241 GATTACAGGCATCTCCAGCGGGGATTTAGGGGCCAGGTGGGGTGGACTTTGACCG 300

Db 387 GATTACAGGCATCTCCAGTGGGACCTAGGGGCCAGGAGTGGACTTTGACCCA 446

Qy 301 TCCACACCCCAAGCTCATGCCCTT 326

Db 447 TCCACCCCAAGCTCATGCCCTT 472

#### RESULT 15

AZ858490/c

LOCUS

DEFINITION

2M0163L10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC2M0163L10 R, DNA sequence.

ACCESSION

AZ858490

VERSION

AZ858490.1

KEYWORDS

GSS.

SOURCE

house musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 658)

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

#### JOURNAL

COMMENT

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: dduun@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0163 row: L column: 10

Seq primer: CACACAGCAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 658.

#### FEATURES

source

1. 658

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone\_lib="UUGC2M0163L10"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adaptor DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 [gi|4732114|gb|AF129072.1], a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 140 a 173 c 177 g 168 t

Query Match 15.2%; Score 272.4; DB 17; Length 658;

Best Local Similarity 86.7%; Pred. No. 6.9e-40;

Matches 300; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1 ATGGGACGCTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCGGGCT 60

Db 582 ATGGGACGCTGAGCAGCCTTATCTCAGGCCACAGTTCACAGCAAGCACTGCGGGCT 523

Qy 61 TCGCAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAACCGGTATTCGACGG 120

Db 522 TCACAGTACAAGCTGCGCAAGTCTCCACCTCAAGAAGCTCAATCGATCTCAGATGG 463

Qy 121 CTGCTGAGTTGGCTTCTCCAGAGCTCCGGTCACGCAAGTCCAGCTCCAAATGGC 180

Db 462 CTGCTGAGTTGGCTTCTCCAGAGCTCCGGTCGGAAGTCAAGTTCGAAATGGGA 403

Qy 181 AAGAGCGAAGACTTCTTCTACATCAAGTTCAGCAGAAAGCCGGGCTCCCATCACCCA 240

Db 402 AAGAGCGAAGACTTCTTCTACATCAAGTTCAGCAGAAAGCCGGGCTCCCATCACCCA 343

Qy 241 GATTACAGGCATCTCCAGCGGGGATTTAGGGGCCAGGTGGGGTGGACTTTGACCG 300

Db 342 GACTACAGCCCTGTCCAGTGGGACATAGGGGTTCAGACCGGAGTAGATTTTCATCCA 283

Qy 301 TCCACACCCCAAGCTCATGCCCTTCTCCCAATCAGCTAGAAATGG 346

Db 282 GCCACCCCAAGCTCATGCCCTTCTCCCAATCAGCTAGAAATGG 237

Mon Jun 16 08:15:58 2003

us-09-513-888c-3.rst

Page 11

Search completed: June 15, 2003, 03:41:58  
Job time : 2494.59 secs

---

**THIS PAGE BLANK (USPTO)**